Arias, Jonas E.; Rubio-Ramírez, Juan Francisco; Shin, Minchul; Waggoner, Daniel F.

Working Paper

Inference based on time-varying SVARs identified with sign restrictions

Working Paper, No. 2024-4

Provided in Cooperation with:
Federal Reserve Bank of Atlanta

Suggested Citation: Arias, Jonas E.; Rubio-Ramírez, Juan Francisco; Shin, Minchul; Waggoner, Daniel F. (2024) : Inference based on time-varying SVARs identified with sign restrictions, Working Paper, No. 2024-4, Federal Reserve Bank of Atlanta, Atlanta, GA, https://doi.org/10.29338/wp2024-04

This Version is available at:
https://hdl.handle.net/10419/300458

FEDERAL RESERVE BANK *of* ATLANTA **WORKING PAPER SERIES**

Inference Based On Time-Varying SVARs Identified with Sign Restrictions

Jonas E. Arias, Juan F. Rubio-Ramírez,
Minchul Shin, and Daniel F. Waggoner

Working Paper 2024-4
March 2024

Abstract: We propose an approach for Bayesian inference in time-varying structural vector autoregressions (SVARs) identified with sign restrictions. The linchpin of our approach is a class of rotation-invariant time-varying SVARs in which the prior and posterior densities of any sequence of structural parameters belonging to the class are invariant to orthogonal transformations of the sequence. Our methodology is new to the literature. In contrast to existing algorithms for inference based on sign restrictions, our algorithm is the first to draw from a uniform distribution over the sequences of orthogonal matrices given the reduced-form parameters. We illustrate our procedure for inference by analyzing the role played by monetary policy during the latest inflation surge.

JEL classification: C11, C51, E52, E58

Key words: time-varying parameters, structural vector autoregressions, identification

https://doi.org/10.29338/wp2024-04

The authors thank Mark Bognanni, Frank Schorfheide, Christian Wolf, and Jonathan Wright for helpful comments. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Atlanta, the Federal Reserve Bank of Philadelphia, or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Please address questions regarding content to Juan Rubio-Ramírez (corresponding author), Emory University, Economics Department, Emory University, Rich Memorial Building, Room 306, Atlanta, GA 30322-2240, juan.rubio-ramirez@emory.edu.
Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's website at www.frbatlanta.org. Click "Publications" and then "Working Papers." To receive e-mail notifications about new papers, use frbatlanta.org/forms/subscribe.

1 Introduction

Structural vector autoregressions (SVARs) featuring time-varying parameters are commonly used in empirical macroeconomics to study a wide range of classical questions such as the economic consequences of policy shocks, oil shocks, the feedback between financial markets and economic activity, and the role played by monetary policy during the Great Inflation.[1]

This paper contributes to this literature by developing a theory of identification and algorithms for Bayesian inference in set-identified time-varying SVARs. Our approach preserves the virtues that popularized set-identified constant parameters SVARs while allowing for time variation in the structural parameters as in the seminal work of Primiceri (2005). The latter is critical to capture changes in the transmission mechanism of how variables respond to structural shocks. Our approach builds on the class of time-varying SVARs defined by Bognanni (2018). The class has two ingredients: a measurement equation linking observed variables to a sequence of time-dependent structural parameters and a prior over such sequence. Notably, the prior and posterior densities over sequences of structural parameters of any model belonging to the class are invariant to orthogonal transformations of the sequences. Accordingly, we will refer to the members of this class as rotation-invariant time-varying SVARs. To solve this identification problem, we will use traditional sign restrictions typically inspired by economic theory and institutional knowledge.

In parallel to the theory of identification in constant parameters SVARs and following Rothenberg (1971), we define observational equivalence in time-varying SVARs models and show that orthogonal transformations of a sequence of structural parameters produce observationally equivalent sequences of structural parameters. This result allows us to trace another parallel with the constant parameters case. More specifically, expanding the work of Uhlig (2005) and Rubio-Ramírez, Waggoner and Zha (2010), we can write sequences of structural parameters as sequences of orthogonal reduced-form parameters and define an invertible mapping between them. With this mapping in hand, we demonstrate that any prior over the time-varying reduced-form parameters combined with an independent

[1] See for example, Primiceri (2005); Sims and Zha (2006b); Baumeister and Peersman (2013); Galí and Gambetti (2015); Amir-Ahmadi, Matthes and Wang (2016); Brunnermeier et al. (2021); Hubrich and Waggoner (2022); Aastveit, Furlanetto and Loria (2023).

uniform prior over the sequences of orthogonal matrices produces an element of the class of rotation-invariant time-varying SVARs models. This is an important if and only if result that delineates the scope of our methodology: commonly used modeling choices for the time-varying reduced-form parameters can be easily adapted for structural analysis—e.g., Primiceri (2005) and the dynamic linear models with discounted Wishart stochastic volatility inspired by Uhlig (1994) and documented in Prado and West (2010) and Bognanni (2018). Next, we describe a particular prior over the time-varying reduced-form parameters exploiting Archakov and Hansen's (2021) novel parametrization of correlation matrices. Hence, we refer to it as Random Correlations prior, demonstrating that it defines an element of the class that we will call Random Correlations SVAR (RC-SVAR). Our rationale for using the RC-SVAR is motivated by the insights in Giannone, Lenza and Primiceri (2015) pointing out that a natural way to determine the impact of priors is to assess their implied out-of-sample forecasting performance. Arias, Rubio-Ramírez and Shin (2023) show that for the empirical applications typically considered, the Random Correlations prior generally implies a higher log-predictive score than most alternative models. Our algorithm will be described in terms of RC-SVARs. However, it could be used for any element in the class of rotation-invariant time-varying SVARs induced by any of the alternative priors over the time-varying reduced-form parameters mentioned above.

We introduce a Gibbs Sampler algorithm to draw from the posterior distribution of the Random Correlations SVAR conditional on the sign restrictions. The algorithm adapts the particle Gibbs with ancestor sampling developed by Lindsten, Jordan and Schon (2014) to sequentially draw from the posterior distribution over the sequences of structural parameters conditional on the sign restrictions. Importantly, existing methods consider a uniform distribution over the sequences of orthogonal matrices conditional on the reduced-form parameters and the sign restrictions (e.g., Baumeister and Peersman, 2013; Bognanni, 2018; Debortoli, Galí and Gambetti, 2020). Although this approach makes posterior sampling straightforward, the priors over time-varying reduced-form parameters and the sequences of orthogonal matrices are not independent. Consequently, this method generates draws of structural parameters that do not belong to the class of rotation-invariant time-varying SVARs.

We illustrate our methods by analyzing the monetary policy tightening cycle that began on March 16, 2022. Since lift-off, policy discussions have revolved around the effects of interest rate increases on economic activity and inflation. In particular, there has been an ongoing debate regarding how much interest rates must increase to achieve this objective. As Powell (2023) recently noted, doing too little or too much could cause unnecessary harm to the economy. Motivated by this discussion, we use our methodology to tackle three questions: (i) How did the Federal Reserve respond to the state of the economy during the current policy tightening cycle? (ii) How does the Federal Reserve's performance during the tightening cycle compare with more dovish or hawkish monetary policy stances? and (iii) Was the Federal Reserve behind the curve? And, if so, at what cost?

As will become apparent, our methodology provides a new framework for answering these questions because it can estimate how the Fed reacted to economic conditions, allowing for the reaction function and the remaining parts of the economy to change over time. This is important for several reasons. Not only has the Federal Reserve adopted different operating procedures such as interest rates and non-borrowed reserves targeting, but also the reaction function may have differed over time within operating procedures (see, e.g., Clarida, Gali and Gertler, 2000). In addition, certain structural relations in the economy may have changed. Our econometric approach provides a helpful setting to discipline the estimation of the reaction function with a relatively small number of novel time-varying sign restrictions. This setup allows us to change the identification restrictions across the sample.

We use our estimates to decompose the quarterly average level of the federal funds rate for each quarter from 2022Q2 until 2023Q2 into three components: the predictable component, the unpredictable component that can be attributed to non-monetary policy shocks (such as demand or supply shocks) hitting the economy; and the unpredictable component that can be attributed to monetary policy shocks. Our estimates suggest that about two thirds of the unpredictable component was a response to non-monetary policy shocks. The response to non-monetary policy shocks is due to the systematic part of the monetary policy. Both systematic monetary policy and monetary policy shocks increased the fed funds rate above the predicted rate in 2022Q2.

To shed light on the second question, we replay history under two counterfactual simula-

tions that we label Dovish Fed and Hawkish Fed. These simulations allow us to determine what would have happened if the Fed had been more or less aggressive relative to the reaction function estimated by our model. In the Dovish (Hawkish) Fed counterfactual, we modify the reaction function so that the response of the federal funds rate to inflation is half (twice) as large as the one in the estimated reaction function for the first quarter of 2022. Focusing on the posterior medians, through the lens of our model, under the Dovish Fed counterfactual, the economy would have marginally overheated, and inflation would have run persistently above 5 percent. Under the Hawkish Fed counterfactual, inflation would have quickly decreased at a small cost in terms of economic activity: real GDP in the second quarter of 2023 would have been about 0.7 percent lower than in the data. Even so, when looking at the level of output at risk, the lower envelope of the 68 percent probability bands shows that the cost in terms of output could have been as large as 3.1 percent.

Turning to the third question, our model estimates provide support to the view that the Federal Reserve was behind the curve (see, e.g., Summers, 2021b) in 2021. Notwithstanding, we also find that the delay in increasing the federal funds rate was not the main driver of the surge in inflation during 2021. Non-monetary shocks explain the unexpected increase in inflation during this time.

The structure of the paper is as follows. Section 2 describes a class of rotation-invariant time-varying SVARs. Section 3 maps such class to orthogonal reduced-form models. Section 4 describes the reduced-form model we will use to illustrate our methodology. Section 5 presents algorithms for inference in rotation-invariant time-varying models identified with sign restrictions. Section 6 studies the role played by monetary policy during the latest tightening cycle. Finally, Section 7 concludes.

2 Rotation-Invariant Time-Varying Structural Models

Consider the general class of time-varying SVARs models defined by Bognanni (2018). Two elements define the class. First, a measurement equation:

$$\mathbf{y}_t' \mathbf{A}_t = \mathbf{x}_t' \mathbf{F}_t + \boldsymbol{\varepsilon}_t' \text{ with } \boldsymbol{\varepsilon}_t \sim \mathrm{N}(\mathbf{0}_n, \mathbf{I}_n) \text{ for } 1 \leq t \leq T, \tag{1}$$

where $\mathbf{y}_t$ is an $n \times 1$ vector of endogenous variables, $\mathbf{x}_t = \begin{bmatrix} 1 & \mathbf{y}'_{t-1} & \cdots & \mathbf{y}'_{t-p} \end{bmatrix}'$ is an $m \times 1$ vector featuring a constant and lag endogenous variables with $m = np + 1$, $\boldsymbol{\varepsilon}_t$ is an $n \times 1$ vector of orthogonal structural shocks, p is the lag length, T is the sample size, and the initial conditions, $(\mathbf{y}_0, \cdots, \mathbf{y}_{1-p})$, are known. The vector $\boldsymbol{\varepsilon}_t$, conditional on time $t-1$ information, is Gaussian with mean zero and covariance matrix $\mathbf{I}_n$, the $n \times n$ identity matrix. The $n \times n$ matrix $\mathbf{A}_t$, which must be invertible, and the $m \times n$ matrix $\mathbf{F}_t$ are the time-varying structural parameters.

The second element is the law of motion for the time-varying structural parameters:

$$p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}) = \prod_{t=1}^T p_S\left(\mathbf{A}_t, \mathbf{F}_t \mid (\mathbf{A}_j, \mathbf{F}_j)_{j=1}^{t-1}, \boldsymbol{\phi}\right), \tag{2}$$

where $\boldsymbol{\phi}$ denotes the constant parameters controlling the evolution of the time-varying structural parameters.[2] We are going to assume that all the elements of the class share the same measurement equation. Thus, different elements of the class are characterized by a different law of motion $p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi})$. The theory we develop is valid for the general law of motion given by Equation (2). To simplify the exposition, all the examples in this paper will assume that the law of motion is Markov, which means that $p_S\left(\mathbf{A}_t, \mathbf{F}_t \mid (\mathbf{A}_j, \mathbf{F}_j)_{j=1}^{t-1}, \boldsymbol{\phi}\right) = p_S\left(\mathbf{A}_t, \mathbf{F}_t \mid \mathbf{A}_{t-1}, \mathbf{F}_{t-1}, \boldsymbol{\phi}\right)$. We assume that the law of motion satisfies:

$$p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}) = p_S((\mathbf{A}_t \mathbf{Q}_t, \mathbf{F}_t \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi}), \tag{3}$$

for every sequence of orthogonal matrices $(\mathbf{Q}_t)_{t=1}^T \in \mathcal{O}_n^T$.[3] This assumption implies that if $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ is any permissible sequence of time-varying structural parameters and $(\mathbf{Q}_t)_{t=1}^T$ is any sequence of orthogonal matrices, then $(\mathbf{A}_t \mathbf{Q}_t, \mathbf{F}_t \mathbf{Q}_t)_{t=1}^T$ is a permissible sequence of time-varying structural parameters such that Equation (3) holds.[4,5]

[2] We follow the convention that $p\left(\mathbf{A}_1, \mathbf{F}_1 \mid (\mathbf{A}_j, \mathbf{F}_j)_{j=1}^0, \boldsymbol{\phi}\right) = p\left(\mathbf{A}_1, \mathbf{F}_1 \mid \boldsymbol{\phi}\right)$.

[3] The notation $\mathcal{O}_n$ denotes the set of all $n \times n$ orthogonal matrices and $\mathcal{O}_n^T = \prod_{t=1}^T \mathcal{O}_n$ the set of all sequences of $n \times n$ orthogonal matrices of length T.

[4] The sequence $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ is permissible if and only if $p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}) > 0$.

[5] When possible, we will avoid the use of the word *sequence* to economize language.

The likelihood of the model can be written as:

$$p((\mathbf{y}_t)_{t=1}^T \mid (\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}) = \prod_{t=1}^T p(\mathbf{y}_t \mid \mathbf{x}_t, \mathbf{A}_t, \mathbf{F}_t). \tag{4}$$

Notice that the likelihood does not depend on $\boldsymbol{\phi}$. Given our assumption about the shocks, the distribution of $\mathbf{y}_t$, conditional on $(\mathbf{x}_t, \mathbf{A}_t, \mathbf{F}_t)$, is Gaussian with mean $\mathbf{x}_t' \mathbf{F}_t \mathbf{A}_t^{-1}$ and variance $(\mathbf{A}_t \mathbf{A}_t')^{-1}$. In particular, $p(\mathbf{y}_t \mid \mathbf{x}_t, \mathbf{A}_t, \mathbf{F}_t)$ can be easily computed. The parameters of this class of time-varying SVAR models are $((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi})$. Hence, the law of motion given by Equation (2) can be interpreted as the prior over the time-varying structural parameters, conditional on $\boldsymbol{\phi}$.[6] The prior can be completed by specifying a marginal prior over the constant parameters, which we will denote $p(\boldsymbol{\phi} \mid \boldsymbol{\psi})$, where $\boldsymbol{\psi}$ are some fixed hyperparameters.[7] Following Rothenberg (1971), $((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi})$ and $((\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T, \tilde{\boldsymbol{\phi}})$ will be observationally equivalent if and only if the likelihoods are equal for almost all $(\mathbf{y}_t)_{t=1}^T \in \mathbb{R}^{nT}$. Because the likelihood does not depend on $\boldsymbol{\phi}$, it makes sense to talk about $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ and $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T$ being observationally equivalent. The following proposition, which is a reminiscence of the constant parameters case, gives a necessary and sufficient condition for the observational equivalence of $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ and $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T$.

Proposition 1. *The time-varying structural parameters* $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ *and* $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T$ *are observationally equivalent if and only if there exists* $(\mathbf{Q}_t)_{t=1}^T \in \mathcal{O}_n^T$ *such that* $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T = (\mathbf{A}_t \mathbf{Q}_t, \mathbf{F}_t \mathbf{Q}_t)_{t=1}^T$.

Proof. See Appendix A. □

A similar (but not identical) proposition can be found in Bognanni (2018). In light of Proposition 1, we can interpret the restriction given by Equation (3) as the necessary and sufficient condition that forces the prior over the time-varying structural parameters, to be equal over observationally equivalent time-varying structural parameters. This restriction ensures that the prior over the time-varying structural parameters does not affect identification and only the identification restrictions influence the posterior distribution of observationally

[6] The prior over the time-varying structural parameters is always conditional on $\boldsymbol{\phi}$, hence we will only write prior over the time-varying structural parameters.

[7] At the cost of more complicated notation, we could also consider a prior over $\boldsymbol{\psi}$.

equivalent time-varying structural parameters. For this reason, we use the term *rotation-invariant* time-varying SVARs models to refer to any structural model consistent with Equations (1)-(3). To solve the identification problem, we will impose sign restrictions on either the sequence of time-varying structural parameters or some function of them, such as sequences of time-varying IRFs. These restrictions will be discussed in Section 5.

2.1 Heteroskedastic Structural Shocks

The restrictions in Equation (3) rule out time-varying SVARs with heteroskedastic structural shocks from belonging to the class of rotation-invariant time-varying SVARs models. Such models have the following common specification:

$$\mathbf{y}_t' \mathbf{A} = \mathbf{x}_t' \mathbf{F} + \tilde{\boldsymbol{\varepsilon}}_t' \text{ with } \tilde{\boldsymbol{\varepsilon}}_t \sim \mathrm{N}(\mathbf{0}_n, \boldsymbol{\Psi}_t) \text{ for } 1 \leq t \leq T,$$

where $\boldsymbol{\Psi}_t$ is an $n \times n$ time-varying diagonal matrix with positive diagonal, $\mathbf{A}$ is a $n \times n$ invertible matrix with ones along the diagonal, and $\mathbf{F}$ is an $m \times n$ matrix. These models can be written in terms of the measurement Equation (1) by defining the $n \times n$ matrix $\mathbf{A}_t = \mathbf{A}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}}$ and the $m \times n$ matrix $\mathbf{F}_t = \mathbf{F}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}}$ as the time-varying structural parameters.[8] Lütkepohl and Netšunajev (2017) describe several ways to model the law of motion of $\boldsymbol{\Psi}_t$.[9] We show that the priors over the time-varying structural parameters implied by models with heteroskedastic structural shocks do not satisfy Equation (3).

The restrictions imply that the prior $p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T | \boldsymbol{\phi})$ must be zero unless there exists $\mathbf{A}$, $\mathbf{F}$, and a sequence of diagonal matrices with positive diagonal $(\boldsymbol{\Psi}_t)_{t=1}^T$ such that $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T = (\mathbf{A}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}}, \mathbf{F}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}})_{t=1}^T$. So, for any permissible sequence $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$, $\mathbf{A}_1^{-1}\, \mathbf{A}_t = \boldsymbol{\Psi}_1^{\frac{1}{2}}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}}$ must be diagonal for $2 \leq t \leq T$. Let $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T = (\mathbf{A}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}}, \mathbf{F}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}})_{t=1}^T$ be any permissible sequence and define $(\mathbf{Q}_t)_{t=1}^T$ so that $\mathbf{Q}_1 \in \mathcal{O}_n$ is not diagonal and $\mathbf{Q}_t = \mathbf{I}_n$, for $2 \leq t \leq T$. The sequence $(\mathbf{A}_t\, \mathbf{Q}_t, \mathbf{F}_t\, \mathbf{Q}_t)_{t=1}^T$ cannot be permissible because $(\mathbf{A}_1\, \mathbf{Q}_1)^{-1}(\mathbf{A}_t\, \mathbf{Q}_t) = \mathbf{Q}_1'\, \boldsymbol{\Psi}_1^{\frac{1}{2}}\, \boldsymbol{\Psi}_t^{-\frac{1}{2}}$ is not diagonal for $2 \leq t \leq T$.

[8] For diagonal matrices, $\boldsymbol{\Psi}_t^{-\frac{1}{2}}$ denotes the element by element inverse square root, either positive or negative, of the elements of $\boldsymbol{\Psi}_t$.

[9] There are also alternative ways of normalizing these models other than fixing each element of the diagonal of $\mathbf{A}$ to one. These alternative normalizations can also be written in terms of of Equation (1).

Similar arguments show that the restriction given by Equation (3) excludes the models described in Sentana and Fiorentini (2001); Rigobon (2003); Rigobon and Sack (2003); Lanne and Lütkepohl (2008); Lanne, Lütkepohl and Maciejowska (2010); Brunnermeier et al. (2021). At this point, one might wonder if the class of rotation-invariant time-varying SVARs models is too small to be of practical use. In the next section, we show how to characterize and easily construct models belonging to the class.

The arguments above highlight the usefulness of considering rotation-invariant time-varying SVARs models. While heteroskedasticity of the structural shocks can be exploited for identification, these shocks may not have a meaningful structural economic interpretation (in fact, Herwartz and Lütkepohl, 2014, try to address this issue). In contrast, using sign restrictions—inspired by economic theory or institutional knowledge—on either the structural parameters or some function of the structural parameters, like impulse responses, ensures that the shocks have a meaningful structural economic interpretation. In addition, the approach of identification through heteroskedasticity requires constant impulse responses up to scale, which has been deemed to be a potential Achilles' heel of the approach (see e.g., Brunnermeier et al., 2021).

3 Time-Varying Orthogonal Reduced-Form Models

Proposition 1 implies that the measurement equation described in Section 2 can alternatively be written in terms of what we call the time-varying orthogonal reduced-form parameters. This parameterization is characterized by a sequence of time-varying reduced-form parameters and orthogonal matrices $(\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T$ and the measurement equation can be written as follows:

$$\mathbf{y}_t' = \mathbf{x}_t' \mathbf{B}_t + \boldsymbol{\varepsilon}_t' \mathbf{Q}_t' h(\mathbf{\Sigma_t}) \text{ for } 1 \leq t \leq T, \tag{5}$$

where the $n \times n$ matrix $h(\mathbf{\Sigma})$ is any decomposition of the precision matrix $\mathbf{\Sigma}$ satisfying $h(\mathbf{\Sigma})' h(\mathbf{\Sigma}) = \mathbf{\Sigma}$. We will take h to be the Cholesky decomposition, normalized so that the diagonal is positive, though any differentiable decomposition would do. As in the case of constant parameters SVARs, the orthogonal reduced-form parameters are convenient for

drawing. The orthogonal reduced-form parameters can be turned into structural parameters by exploiting the following mapping from the time-varying structural parameters to the time-varying orthogonal reduced-form parameters:

$$f_h((\mathbf{A}_t,\mathbf{F}_t)_{t=1}^T)=\Big(\underbrace{\mathbf{F}_t\,\mathbf{A}_t^{-1}}_{\mathbf{B}_t},\underbrace{(\mathbf{A}_t\,\mathbf{A}_t')^{-1}}_{\mathbf{\Sigma}_t},\underbrace{h((\mathbf{A}_t\,\mathbf{A}_t')^{-1})\,\mathbf{A}_t}_{\mathbf{Q}_t}\Big)_{t=1}^T. \tag{6}$$

This function is invertible and its inverse is given by:

$$f_h^{-1}((\mathbf{B}_t,\mathbf{\Sigma}_t,\mathbf{Q}_t)_{t=1}^T)=\Big(\underbrace{h(\mathbf{\Sigma}_t)^{-1}\,\mathbf{Q}_t}_{\mathbf{A}_t},\underbrace{\mathbf{B}_t\,\mathbf{A}_t}_{\mathbf{F}_t}\Big)_{t=1}^T. \tag{7}$$

Let $p_{OR}((\mathbf{B}_t,\mathbf{\Sigma}_t,\mathbf{Q}_t)_{t=1}^T\mid\boldsymbol{\phi})$ denote a prior over the time-varying orthogonal reduced-form parameters, conditional on $\boldsymbol{\phi}$.[10] The functions defined by Equations (6) and (7) allow us to transform priors over the time-varying structural parameters into equivalent priors over the time-varying orthogonal reduced-form parameters and vice versa, but we must take into account the volume element of the transformations. By Proposition 1 of Arias, Rubio-Ramírez and Waggoner (2018), the volume element of the mapping given by Equation (6) is $v_{f_h}((\mathbf{A}_t,\mathbf{F}_t)_{t=1}^T)=\prod_{t=1}^T 2^{\frac{n(n+1)}{2}}|\det(\mathbf{A}_t)|^{-(2n+m+1)}$ and the volume element of the mapping given by Equation (7) is $v_{f_h^{-1}}((\mathbf{B}_t,\mathbf{\Sigma}_t,\mathbf{Q}_t)_{t=1}^T)=\prod_{t=1}^T 2^{-\frac{n(n+1)}{2}}|\det(\mathbf{\Sigma}_t)|^{-\frac{2n+m+1}{2}}$. Thus, the prior over the time-varying orthogonal reduced-form parameters induced by p_S is:

$$p_{OR}((\mathbf{B}_t,\mathbf{\Sigma}_t,\mathbf{Q}_t)_{t=1}^T\mid\boldsymbol{\phi})=2^{-\frac{n(n+1)T}{2}}\Big(\prod_{t=1}^T|\det(\mathbf{\Sigma}_t)|\Big)^{-\frac{2n+m+1}{2}}p_S(f_h^{-1}((\mathbf{B}_t,\mathbf{\Sigma}_t,\mathbf{Q}_t)_{t=1}^T)\mid\boldsymbol{\phi}),$$

and the prior over the time-varying structural parameters induced by p_{OR} is:

$$p_S((\mathbf{A}_t,\mathbf{F}_t)_{t=1}^T\mid\boldsymbol{\phi})=2^{\frac{n(n+1)T}{2}}\Big(\prod_{t=1}^T|\det(\mathbf{A}_t)|\Big)^{-(2n+m+1)}p_{OR}(f_h((\mathbf{A}_t,\mathbf{F}_t)_{t=1}^T)\mid\boldsymbol{\phi}).$$

The following proposition shows that Equation (3) translates into a restriction on the prior over the time-varying orthogonal reduced-form parameters.

[10] The prior over the time-varying orthogonal reduced-form parameters is always conditional on $\boldsymbol{\phi}$, hence we will only write prior over the time-varying orthogonal reduced-form parameters.

Proposition 2. *The prior over the time-varying structural parameters satisfies Equation* (3) *if and only if the induced prior over the time-varying orthogonal reduced-form parameters does not depend on* $(\mathbf{Q}_t)_{t=1}^T$.

Proof. See Appendix A. □

The proposition implies that for any prior over the time-varying structural parameters satisfying Equation (3), the induced prior over the time-varying orthogonal reduced-form parameters must be independent over $(\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T$ and $(\mathbf{Q}_t)_{t=1}^T$, and the induced prior over $(\mathbf{Q}_t)_{t=1}^T$ must be uniform with respect to the volume measure over $\mathcal{O}_n^T$.[11] More importantly, it also says that any prior over the time-varying orthogonal reduced-form parameters such that the prior over the orthogonal matrices conditional on the time-varying reduced-form parameters is uniform induces a prior over the structural parameters that satisfies the restriction given by Equation (3).[12] In other words, any prior over the time-varying orthogonal reduced-form parameters that can be written as:

$$p_{OR}((\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi}) = \frac{p_R((\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi})}{\prod_{t=1}^T \int_{\mathcal{O}(n)} 1 d_{\mathcal{O}(n)}\mathbf{Q}}$$

induces a prior over the structural parameters that satisfies the restriction given by Equation (3), where $p_R((\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi})$ denotes the prior over the time-varying reduced-form parameters. Hence, every prior over $(\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T$ corresponds to an element of the class of rotation-invariant time-varying SVARs models, and the prior over the time-varying structural parameters induced by p_R is:

$$p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}) = 2^{\frac{n(n+1)T}{2}} \Big(\prod_{t=1}^T |\det(\mathbf{A}_t)| \Big)^{-(2n+m+1)} \frac{p_R(\pi(f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T)) \mid \boldsymbol{\phi})}{\prod_{t=1}^T \int_{\mathcal{O}(n)} 1 d_{\mathcal{O}(n)}\mathbf{Q}},$$

where $\pi(\cdot)$ denotes the projection of $(\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T$ onto $(\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T$. In the next section, we use the novel parameterization of the correlation matrix described in Archakov and Hansen (2021) to define a prior over the time-varying reduced-form parameters as in Arias, Rubio-

[11] The proposition generalizes Proposition 4 in Bognanni (2018).

[12] Furthermore, it is easy to see that in this setup the prior over the time-varying reduced-form parameters will be Markov if and only if the prior over the time-varying structural parameters is Markov.

Ramírez and Shin (2023). This prior defines an element of the class of rotation-invariant time-varying SVARs models.

The above results depend only on the fact that the volume element does not depend on the sequence of orthogonal matrices. Hence, Proposition 2 will extend to any alternative parameterization of the class, provided that the volume element does not depend on the sequence of orthogonal matrices. For example, this would be the case if the measurement equation were written in terms of impulse responses.

4 Time-Varying Reduced-Form Model

In this section, we describe the prior over the time-varying reduced-form parameters that will be used in the rest of the paper. It is based on a time-varying extension of the parameterization for reduced-form variance-covariance matrices proposed by Archakov and Hansen (2021). Consider the decomposition of the reduced-form variance-covariance matrix given by $\boldsymbol{\Sigma}_t = \mathbf{D}_t \mathbf{C}_t \mathbf{D}_t$, where $\mathbf{D}_t = \text{diag}(\text{diag}(\boldsymbol{\Sigma}_t)^{\frac{1}{2}})$ is the diagonal matrix containing the standard deviations and $\mathbf{C}_t = \mathbf{D}_t^{-1} \boldsymbol{\Sigma}_t \mathbf{D}_t^{-1}$ is the correlation matrix.[13] We can map $\boldsymbol{\Sigma}_t$ to $(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t) \in \mathbb{R}^n \times \mathbb{R}^{n_\gamma}$, where $n_\gamma = n(n-1)/2$, $\boldsymbol{\delta}_t = 2\log(\text{diag}(\mathbf{D}_t))$, and $\boldsymbol{\gamma}_t = \text{vecl}(\log(\mathbf{C}_t))$.[14] Clearly, the mapping $\mathbf{D}_t \to \boldsymbol{\delta}_t$ is invertible and, by Theorem 1 of Archakov and Hansen (2021), the mapping $\mathbf{C}_t \to \boldsymbol{\gamma}_t$ is also invertible. Thus, we can define an invertible function $g^{RC}((\boldsymbol{\Sigma}_t)_{t=1}^T) = (\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T$, and as a consequence, any law of motion, or equivalently any prior, defined over $(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T$ translates into a prior over the time-varying reduced-form parameters. Consider the following law of motion for $(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=2}^T$:

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t, \text{ with } \boldsymbol{\nu}_t \sim \text{N}(\mathbf{0}_{nm}, \mathbf{V}_{\boldsymbol{\beta}}) \text{ and } \boldsymbol{\beta}_t = \text{vec}(\mathbf{B}_t), \tag{8}$$

$$\boldsymbol{\delta}_t = \boldsymbol{\delta}_{t-1} + \boldsymbol{\eta}_t, \text{ with } \boldsymbol{\eta}_t \sim \text{N}(\mathbf{0}_n, \mathbf{V}_{\boldsymbol{\delta}}), \tag{9}$$

$$\boldsymbol{\gamma}_t = \boldsymbol{\gamma}_{t-1} + \boldsymbol{\zeta}_t, \text{ with } \boldsymbol{\zeta}_t \sim \text{N}(\mathbf{0}_{n_\gamma}, \mathbf{V}_{\boldsymbol{\gamma}}), \tag{10}$$

[13] The linear operator $\text{diag}(\cdot)$, when applied to a vector, denotes the diagonal matrix with the vector along the diagonal and, when applied to a square matrix, denotes the diagonal of the matrix. The square root is the element-by-element positive square root.

[14] The linear operator $\text{vecl}(\cdot)$ returns the vectorized strictly lower triangular component of a square matrix. When applied to a vector, the function $\log(\cdot)$ denotes the element-by-element logarithm and, when applied to a square matrix, denotes the matrix logarithm.

where $\mathbf{V}_{\boldsymbol{\beta}}$ is a symmetric definite positive $nm \times nm$ matrix, $\mathbf{V}_{\boldsymbol{\delta}} = \text{diag}(\text{V}_{\delta,1}, \ldots, \text{V}_{\delta,\text{n}})$ is a diagonal definite positive $n \times n$ matrix and $\mathbf{V}_{\boldsymbol{\gamma}} = \text{diag}(\text{V}_{\gamma,1}, \ldots, \text{V}_{\gamma,\text{n}_\gamma})$ is a diagonal definite positive $n_\gamma \times n_\gamma$ matrix. In addition, we assume that $\boldsymbol{\beta}_1 \sim \text{N}(\mathbf{m}_{\boldsymbol{\beta}_1}, \mathbf{V}_{\boldsymbol{\beta}_1})$, where $\mathbf{m}_{\boldsymbol{\beta}_1}$ is a $nm \times 1$ vector and $\mathbf{V}_{\boldsymbol{\beta}_1}$ is a symmetric definite positive $nm \times nm$ matrix; $\boldsymbol{\delta}_1 \sim \text{N}(\mathbf{m}_{\boldsymbol{\delta}_1}, \mathbf{V}_{\boldsymbol{\delta}_1})$, where $\mathbf{m}_{\boldsymbol{\delta}_1}$ is an $n \times 1$ vector and $\mathbf{V}_{\boldsymbol{\delta}_1}$ is an $n \times n$ diagonal matrix with positive diagonal; and $\boldsymbol{\gamma}_1 \sim \text{N}(\mathbf{m}_{\boldsymbol{\gamma}_1}, \mathbf{V}_{\boldsymbol{\gamma}_1})$, where $\mathbf{m}_{\boldsymbol{\gamma}_1}$ is an $n_\gamma \times 1$ vector and $\mathbf{V}_{\boldsymbol{\gamma}_1}$ is an $n_\gamma \times n_\gamma$ diagonal matrix with positive diagonal. It is straightforward to see that the constant parameters of the model are:

$$\boldsymbol{\phi}^{RC} = (\text{vech}(\mathbf{V}_{\boldsymbol{\beta}}), \text{diag}(\mathbf{V}_{\boldsymbol{\delta}}), \text{diag}(\mathbf{V}_{\boldsymbol{\gamma}}), \mathbf{m}_{\boldsymbol{\beta}_1}, \text{vech}(\mathbf{V}_{\boldsymbol{\beta}_1}), \mathbf{m}_{\boldsymbol{\delta}_1}, \text{diag}(\mathbf{V}_{\boldsymbol{\delta}_1}), \mathbf{m}_{\boldsymbol{\gamma}_1}, \text{diag}(\mathbf{V}_{\boldsymbol{\gamma}_1}))\,,$$

which is a vector of dimension $n^{RC} = \frac{nm(nm+1)}{2} + n + n_\gamma + nm + \frac{nm(nm+1)}{2} + 2n + 2n_\gamma$. Appendix B describes the prior over $\boldsymbol{\phi}^{RC}$. We denote the above prior over $(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T$ by $p^{RC}((\mathbf{B}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) p^{RC}((\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC})$, which we call the Random Correlations prior. This prior, via the function g^{RC} and the identity mapping, induces a prior over the time-varying reduced-form parameters, denoted by $p_R^{RC}((\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC})$, of the form:

$$p_R^{RC}((\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) = p^{RC}((\mathbf{B}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) v_{g^{RC}}((\boldsymbol{\Sigma}_t)_{t=1}^T) p^{RC}(g^{RC}((\boldsymbol{\Sigma}_t)_{t=1}^T) \mid \boldsymbol{\phi}^{RC}), \quad (11)$$

where the volume element $v_{g^{RC}}((\boldsymbol{\Sigma}_t)_{t=1}^T)$ can be computed numerically.[15]

If we combine this prior over the time-varying reduced-form parameters with the uniform prior over $(\mathbf{Q}_t)_{t=1}^T$, the results of Section 3 imply that the induced prior over the time-varying structural parameters will satisfy the restrictions given by Equation (3). In particular, we are going to consider a prior over the time-varying structural parameters, denoted by $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC})$, equal to:

$$p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) = 2^{\frac{n(n+1)T}{2}} \left(\prod_{t=1}^T |\det(\mathbf{A}_t)| \right)^{-(2n+m+1)} \frac{p_R^{RC}(\pi(f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T))) \mid \boldsymbol{\phi}^{RC})}{\prod_{t=1}^T \int_{\mathcal{O}(n)} 1 d_{\mathcal{O}(n)}\mathbf{Q}}, \quad (12)$$

[15] Because $\boldsymbol{\Sigma}_t$ is symmetric, there is an implicit linear restriction on $\boldsymbol{\Sigma}_t$. To directly compute the volume element associated with the function $g^{RC}()$, restricted to symmetric matrices, would require Theorem 3, as opposed to the simpler Theorem 2, of Arias, Rubio-Ramírez and Waggoner (2018). Either approach will give identical answers, though using Theorem 3 may be more efficient numerically.

where $\pi(\cdot)$ denotes the projection of $(\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T$ onto $(\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T$. This prior corresponds to an element of the class of rotation-invariant time-varying SVARs models that we will call Random Correlations SVAR (RC-SVAR). Together with a prior over the constant parameters $p^{RC}(\phi^{RC})$ and the likelihood in Equation (4), our prior over the time-varying structural parameters implies a posterior over the time-varying structural parameters that we label $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$.

4.1 Alternative Time-Varying Reduced-Form Models

Although our algorithms will be written in terms of the RC-SVAR, we could write them in terms of other class members. For example, we could consider the member corresponding to the prior over the time-varying reduced-form parameters in Primiceri (2005), which relies on a decomposition of the reduced-form variance-covariance matrix given by $\mathbf{\Sigma}_t = \mathbf{\Xi}_t \mathbf{\Omega}_t \mathbf{\Omega}_t' (\mathbf{\Xi}_t')^{-1}$, where $\mathbf{\Omega}_t$ is a positive diagonal matrix and $\mathbf{\Xi}_t$ is a lower triangular matrix with ones along the diagonal. The prior over $(\mathbf{B}_t, \mathbf{\Xi}_t, \mathbf{\Omega}_t)_{t=1}^T$ defined in Primiceri (2005), which we label $p^P((\mathbf{B}_t)_{t=1}^T \mid \boldsymbol{\phi}^P) p^P((\mathbf{\Xi}_t, \mathbf{\Omega}_t)_{t=1}^T \mid \boldsymbol{\phi}^P)$, and the invertible function $g^P((\mathbf{\Sigma}_t)_{t=1}^T) = (\mathbf{\Xi}_t, \mathbf{\Omega}_t)_{t=1}^T$ induce a prior over the time-varying reduced-form parameters denoted by $p_R^P((\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi}^P)$, of the form:

$$p_R^P((\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi}^P) = p^P((\mathbf{B}_t)_{t=1}^T \mid \boldsymbol{\phi}^P) v_{g^P}((\mathbf{\Sigma}_t)_{t=1}^T) p^P(g^P((\mathbf{\Sigma}_t)_{t=1}^T) \mid \boldsymbol{\phi}^P),$$

where the volume element $v_{g^P}((\mathbf{\Sigma}_t)_{t=1}^T)$ can be computed numerically. This prior over the time-varying reduced-form parameters along with the uniform prior over the sequences of orthogonal matrices induces an alternative prior over the time-varying structural parameters satisfying the restrictions given by Equation (3). We denote such prior by $p_S^P((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^P)$ and it is equal to:

$$p_S^P((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^P) = 2^{\frac{n(n+1)T}{2}} \Big(\prod_{t=1}^T |\det(\mathbf{A}_t)| \Big)^{-(2n+m+1)} \frac{p_R^P(\pi(f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T))) \mid \boldsymbol{\phi}^P)}{\prod_{t=1}^T \int_{\mathcal{O}(n)} 1 d_{\mathcal{O}(n)} \mathbf{Q}}.$$

Together with a prior over the constant parameters $p^P(\phi^P)$ and the likelihood in Equation (4), this prior over the time-varying structural parameters implies a posterior over the time-

varying structural parameters that we label $p_S^P((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^P \mid (\mathbf{y}_t)_{t=1}^T)$. Since the order of the variables matters in this framework (see Bognanni, 2018), there are in fact $n!$ different elements of the class and posteriors, where n is the number of variables.

The fact that Primiceri's (2005) approach has the unappealing feature of being order-dependent has motivated a quest for order-independent approaches. In particular, as in our baseline prior over the reduced-form parameters, the prior over the time-varying reduced-form parameters defined in Bognanni (2018) is order invariant. This approach relies on the discounted Wishart stochastic volatility model to directly define a prior over the time-varying reduced-form parameters, that we label $p_R^{DW}((\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{DW})$. When combined with the uniform prior over the sequences of orthogonal matrices, it implies a prior over the time-varying structural parameters, denoted by $p_S^{DW}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{DW})$, where:

$$p_S^{DW}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{DW}) = 2^{\frac{n(n+1)T}{2}} \Big(\prod_{t=1}^T |\det(\mathbf{A}_t)| \Big)^{-(2n+m+1)} \frac{p_R^{DW}(\pi(f_h(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) \mid \boldsymbol{\phi}^{DW})}{\prod_{t=1}^T \int_{\mathcal{O}(n)} 1 d_{\mathcal{O}(n)} \mathbf{Q}}.$$

Together with a prior over the constant parameters $p^{DW}(\phi^{DW})$ and the likelihood in Equation (4), this prior over the time-varying structural parameters implies a posterior over the time-varying structural parameters that we label $p_S^{DW}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{DW} \mid (\mathbf{y}_t)_{t=1}^T)$.

Our rationale for using the Random Correlations prior is motivated by the insights in Giannone, Lenza and Primiceri (2015) pointing out that a natural way to determine the impact of priors is to assess their implied out-of-sample forecasting performance. For the type of empirical applications considered below, the Random Correlations prior generally implies a higher log-predictive score than most of the orderings of the variables in Primiceri (2005)–the workhorse model in the literature. Although not affected by the order of the variables, the Drifting SVAR proposed by Bognanni (2018) specification is very restrictive and the implied marginal likelihood significantly underperforms our approach.

Finally, as explained in Arias, Rubio-Ramírez and Shin (2023), we could also have followed Asai and McAleer (2009) and directly imposed a Wishart process-based prior on the dynamics of the correlation matrix without the need to parameterize it. While the empirical performance is similar across both procedures, we favor the approach presented here because it preserves the spirit of the random walk modeling in Primiceri (2005). Chan, Koop and Yu (2021) also

describe a prior over the time-varying reduced-form parameters that could be used.

5 Algorithms

This section presents algorithms to draw from the posterior distribution of rotation-invariant time-varying SVARs models conditional on sign restrictions. To facilitate the exposition we will use the RC-SVAR, but our algorithms could be easily adapted for any model of the class of rotation-invariant time-varying SVARs models and, in particular, to the two alternative models described in Section 4.1.

We present three algorithms. The first algorithm is straightforward but infeasible for the sample sizes commonly encountered in empirical macroeconomics unless the identifying sign restrictions are limited to a small number of periods. The second algorithm is typically employed by current papers. This algorithm is feasible, but we show that it unfortunately does not draw from the desired target posterior distribution. The third algorithm draws from the desired target posterior distribution subject to sign restrictions on an arbitrary number of periods. Thus, it overcomes the limitations of the two algorithms mentioned above.

Let $\boldsymbol{S}_t(\mathbf{A}_t, \mathbf{F}_t)$ be any continuous function whose range is $\mathbb{R}^{s_t}$, where s_t is the number of sign restrictions at time t. We will consider sign restrictions of the form $\boldsymbol{S}_t(\mathbf{A}_t, \mathbf{F}_t) > 0$. For instance, the function $\boldsymbol{S}_t()$ could be a collection of impulse responses of various variables to various shocks at various horizons. While we allow the function to change across t, we will suppress the subscript when the function $\boldsymbol{S}_t()$ is identical across t. These are very general types of sign restrictions, but they do not allow for restrictions to combine information across different t. The advantage of this is that we can determine if the sign restrictions are satisfied independently across t. Allowing for time-varying sign restrictions is another important contribution of our methodology. For example, researchers interested in identifying monetary policy rules (see, e.g., Christiano, Eichenbaum and Evans, 1996; Arias, Caldara and Rubio-Ramírez, 2019) typically choose the federal funds rate to be the monetary policy instrument for the entire sample under analysis. This assumption is questionable because the funds rate has not always been the policy instrument: the Federal Reserve targeted non-borrowed reserves during the early years of Chairman Paul Volcker's tenure. As we

demonstrate in Section 6, time-varying sign restrictions would allow these researchers to tackle the challenge by relaxing the assumption that the federal funds rate is the policy instrument throughout the entire sample.

We organize the rest of this section in three parts. The first part presents the simple but limited algorithm. This algorithm will also be useful to establish a connection between our efficient algorithm and an alternative importance sampling algorithm that could be applied in some special cases. The second part describes the current algorithms and shows that they do not draw from the desired target posterior distribution. The third part develops an efficient algorithm based on the Gibbs sampler.

5.1 A Simple Algorithm

Because $\boldsymbol{S}_t()$ is continuous, the set of all sequences of time-varying structural parameters satisfying the sign restrictions will be open in the set of all sequences of time-varying structural parameters. If at least one sequence satisfies the sign restrictions, then the set of sequences satisfying the sign restrictions will be of positive Lebesgue measure in the set of all sequences. This justifies algorithms of the following type to accomplish our objective.

Algorithm 1. *The following algorithm draws from* $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ *conditional on the sign restrictions.*

1. *Let* $M > 0$ *and* $I > 1$ *and set* $i = 1$.
2. *Draw* $((\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i)_{t=1}^T, \boldsymbol{\phi}^{i,RC})$ *from the* $p_R^{RC}((\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ *distribution.*
3. *Draw sequences* $(\mathbf{Q}_t^{i,m})_{t=1}^T$ *independently from the uniform distribution over* $\mathcal{O}_n^T$ *for* $1 \leq m \leq M$.
4. *Let* $(\mathbf{A}_t^{i,m}, \mathbf{F}_t^{i,m})_{t=1}^T = f_h^{-1}((\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i, \mathbf{Q}_t^{i,m})_{t=1}^T)$ *for* $1 \leq m \leq M$.
5. *If* $i < I$, *let* $i = i + 1$ *and return to Step 2.*
6. *Keep* $((\mathbf{A}_t^{i,m}, \mathbf{F}_t^{i,m})_{t=1}^T, \boldsymbol{\phi}^{i,RC})$ *for* $1 \leq i \leq I$ *and* $1 \leq m \leq M$, *if* $\boldsymbol{S}_t(\mathbf{A}_t^{i,m}, \mathbf{F}_t^{i,m}) > \mathbf{0}$ *for* $1 \leq t \leq T$.

Step 2 is implemented by drawing $((\mathbf{B}_t^i, \boldsymbol{\delta}_t^i, \boldsymbol{\gamma}_t^i)_{t=1}^T, \boldsymbol{\phi}^{i,RC})$ from the $p^{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ density and setting $(\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i)_{t=1}^T$ equal to $(g^{RC})^{-1}((\mathbf{B}_t^i, \boldsymbol{\delta}_t^i, \boldsymbol{\gamma}_t^i)_{t=1}^T)$. As mentioned in Section 4.1, using a different prior over the $(\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i)_{t=1}^T$ allows this algorithm to be used for rotation-invariant time-varying SVARs models. One could use the prior over the time-varying reduced-form parameters in either Primiceri (2005) or Bognanni (2018) by simply changing Step 2 to draw from the distributions over the time-varying reduced-form parameters defined in those papers.

Algorithm 1 is very easy to implement. Still, unless the identifying sign restrictions are limited to a few periods, the number of sequences of orthogonal matrices required to get sufficient draws that satisfy the sign restrictions is computationally infeasibly large. To see this, notice that, for every i, the probability that $f_h^{-1}(\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i, \mathbf{Q}_t^{i,m})$ satisfies the restrictions is less than one for $1 \leq t \leq T$. Hence, the probability that $f_h^{-1}((\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i, \mathbf{Q}_t^{i,m})_{t=1}^T)$ satisfies the restrictions converges to zero as T goes to infinity.

5.2 Current Algorithms

Because of the infeasibility of using Algorithm 1 in most applications of interest, current algorithms (e.g., Baumeister and Peersman, 2013; Bognanni, 2018; Debortoli, Galí and Gambetti, 2020) modify Step 3 to draw $(\mathbf{Q}_t^{i,m})_{m=1}^M$ independently from the uniform distribution over $\mathbb{O}_n(\mathbf{B}_t, \boldsymbol{\Sigma}_t) = \left\{\mathbf{Q}_t \in \mathcal{O}_n : \boldsymbol{S}_t(f_h^{-1}(\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)) > \mathbf{0}\right\}$ for $1 \leq m \leq M$ and $1 \leq t \leq T$. This is implemented by drawing $\mathbf{Q}_t$ independently from the uniform distribution over $\mathcal{O}_n$ until one obtains a draw such that $\mathbf{Q}_t \in \mathbb{O}_n(\mathbf{B}_t, \boldsymbol{\Sigma}_t)$, which simplifies computation by generating $\mathbf{Q}_t \in \mathbb{O}_n(\mathbf{B}_t, \boldsymbol{\Sigma}_t)$ period by period.[16] The current algorithms can be written as follows:

Algorithm 2. *The following algorithm draws from Equation (13) conditional on the sign restrictions.*

1. *Let $I > 1$ and set $i = 1$.*

2. *Draw $((\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i)_{t=1}^T, \boldsymbol{\phi}^{i,RC})$ from the $p_R^{RC}((\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ distribution.*

[16] To simplify the exposition, we are assuming that $\mathbb{O}_n(\mathbf{B}_t, \boldsymbol{\Sigma}_t) \neq \varnothing$ for all $(\mathbf{B}_t, \boldsymbol{\Sigma}_t)$ for $1 \leq t \leq T$. The same argument is valid otherwise, although the discussion is more tedious.

3. *Draw* $\mathbf{Q}_t^i$ *uniformly from the set* $\mathbb{O}_n(\mathbf{B}_t^i, \mathbf{\Sigma}_t^i) = \left\{\mathbf{Q}_t \in \mathcal{O}_n : \boldsymbol{S}_t(f_h^{-1}(\mathbf{B}_t^i, \mathbf{\Sigma}_t^i, \mathbf{Q}_t)) > \mathbf{0}\right\}$ *for all* $t = 1, 2, ..., T$. *This step can be done by drawing* $\mathbf{Q}_t$ *from* $\mathcal{O}_n$ *until one that satisfies* $\boldsymbol{S}_t(f_h^{-1}(\mathbf{B}_t^i, \mathbf{\Sigma}_t^i, \mathbf{Q}_t)) > \mathbf{0}$ *is found.*

4. *Let* $(\mathbf{A}_t^i, \mathbf{F}_t^i)_{t=1}^T = f_h^{-1}((\mathbf{B}_t^i, \mathbf{\Sigma}_t^i, \mathbf{Q}_t^i)_{t=1}^T)$.

5. *If* $i < I$, *let* $i = i + 1$ *and return to Step 2.*

The computational benefit comes from the *conditional* nature of the prior distribution for $\mathbf{Q}_t$. However, it has a potentially undesirable implication. For any $(\mathbf{B}_t, \mathbf{\Sigma}_t)$, the current algorithms implicitly define a density $p(\mathbf{Q}_t \mid \mathbf{B}_t, \mathbf{\Sigma}_t)$ with respect to the volume measure over $\mathbb{O}_n(\mathbf{B}_t, \mathbf{\Sigma}_t)$ that is proportional to $1_{\{\mathbf{Q}_t \in \mathbb{O}_n(\mathbf{B}_t, \mathbf{\Sigma}_t)\}}$:

$$p(\mathbf{Q}_t \mid \mathbf{B}_t, \mathbf{\Sigma}_t) = \begin{cases} \frac{1}{\int_{\mathbb{O}_n(\mathbf{B}_t, \mathbf{\Sigma}_t)} 1 d_{\mathcal{O}_n} \mathbf{Q}_t} \text{ if } \mathbf{Q}_t \in \mathbb{O}_n(\mathbf{B}_t, \mathbf{\Sigma}_t) \\ 0 \text{ otherwise} \end{cases}$$

for $1 \leq t \leq T$, which demonstrates that current algorithms define a marginal density for the orthogonal matrices conditional on the reduced-form parameters that is not uniform. To see that, notice that $\int_{\mathbb{O}_n(\mathbf{B}_t, \mathbf{\Sigma}_t)} 1 d_{\mathcal{O}_n} \mathbf{Q}_t$ depends on the reduced-form parameters. Thus, current algorithms draw from:

$$\left(\prod_{t=1}^T |\det(\mathbf{A}_t)|\right)^{-(2n+m+1)} \frac{p_R^{RC}(\pi(f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T))) \mid \boldsymbol{\phi}^{RC})}{\prod_{t=1}^T \int_{\mathbb{O}_n(\pi(f_h(\mathbf{A}_t, \mathbf{F}_t)))} 1 d_{\mathcal{O}_n} \mathbf{Q}_t}, \tag{13}$$

not from the $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ conditional on the sign restrictions. Hence, the prior implied by the current algorithms does not satisfy Proposition 2, and therefore, the implied prior does not impose the same prior probability density on the observationally equivalent path of the structural parameters.

Because Algorithm 2 generates draws from Equation (13) conditional on the sign restrictions and our objective is to draw from the $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ conditional on the

sign restrictions, one could use importance sampling weights proportional to:

$$\frac{\prod_{t=1}^{T}\int_{\mathbb{O}_n(\pi(f_h(\mathbf{A}_t,\mathbf{F}_t)))} 1 d_{\mathcal{O}_n}\mathbf{Q}_t}{\prod_{t=1}^{T}\int_{\mathbb{O}_n} 1 d_{\mathcal{O}_n}\mathbf{Q}_t}$$

to re-sample the outcome of Algorithm 2. However, as explained in Durbin and Koopman (2012), this is excessively burdensome for the type of values that T takes in empirical applications.

5.3 A Gibbs Sampler Algorithm

In this section, we present a feasible Gibbs sampler that draws from the desired target distribution. The key insight is to draw sequentially rather than simultaneously, significantly improving computational efficiency upon Algorithm 1. Unlike current algorithms, our algorithm generates posterior draws based on the prior distribution that satisfies the conditions in Proposition 2. We present the algorithm's main steps and relegate the details to Appendix C.

Algorithm 3. *The following algorithm draws from* $p_S^{RC}((\mathbf{A}_t,\mathbf{F}_t)_{t=1}^T,\boldsymbol{\phi}^{RC}\mid(\mathbf{y}_t)_{t=1}^T)$ *conditional on the sign restrictions.*

1. *Let* $I>1$*, set* $i=1$*, and initialize* $((\mathbf{A}_t^{i-1})_{t=1}^T,\boldsymbol{\phi}^{i-1,RC})$.
2. *Draw* $(\mathbf{F}_t^i)_{t=1}^T$ *from a distribution that is proportional to* $p_S^{RC}((\mathbf{F}_t)_{t=1}^T\mid(\mathbf{A}_t^{i-1})_{t=1}^T,\boldsymbol{\phi}^{i-1,RC},(\mathbf{y}_t)_{t=1}^T)$.
3. *Draw* $(\mathbf{A}_t^i)_{t=1}^T$ *from* $p_S^{RC}((\mathbf{A}_t)_{t=1}^T\mid(\mathbf{F}_t^i)_{t=1}^T,(\mathbf{A}_t^{i-1})_{t=1}^T,\boldsymbol{\phi}^{i-1,RC},(\mathbf{y}_t)_{t=1}^T)$ *conditional on the sign restrictions.*
4. *Draw* $(\boldsymbol{\phi}^{i,RC})$ *from* $p_S^{RC}(\boldsymbol{\phi}^{RC}\mid(\mathbf{A}_t^i,\mathbf{F}_t^i)_{t=1}^T,(\mathbf{y}_t)_{t=1}^T)$.
5. *If* $i<I$*, let* $i=i+1$ *and return to Step 2.*

Step 1 is straightforward, $(\mathbf{A}_t^0)_{t=1}^T$ can be initialized using a constant parameters SVAR or using the point-wise mean of draws obtained from the RC-SVAR without sign restrictions, and $\boldsymbol{\phi}^{0,RC}$ can be drawn from the prior defined in Appendix B.

The critical steps of this algorithm are Steps 2 and 3 because the conditional posterior distribution of $\mathbf{F}_t$ and $\mathbf{A}_t$ conditional on the sign restrictions are not conditionally Gaussian

any more. To overcome this difficulty, we rely on simulation smoothers based on sequential Monte Carlo methods. Importantly, in the case of Step 3, we use the Particle Gibbs Ancestor Sampling (PGAS) algorithm developed by Lindsten, Jordan and Schon (2014). The PGAS algorithm is a variant of sequential Monte Carlo methods. Importantly, when drawing $(\mathbf{A}_t^i)_{t=1}^T$, the PGAS conditions on the previous draw, $(\boldsymbol{A}_t^{i-1})_{t=1}^T$. This ensures that the algorithm has an invariant property, which is necessary for our algorithm to be valid. We refer the reader to Andrieu, Doucet and Holenstein (2010) and Lindsten, Jordan and Schon (2014) for details.

6 The Current Monetary Policy Tightening Cycle

This section illustrates our methodology by analyzing the monetary policy tightening cycle that began on March 16, 2022. Since lift-off, policy discussions have revolved around the effects of interest rate increases on economic activity and inflation. In particular, there has been an ongoing debate regarding how much interest rates must increase to achieve this objective. As Powell (2023) recently noted, doing too little or too much could cause unnecessary harm to the economy. Motivated by this discussion, we use our methodology to tackle three questions: (i) How did the Federal Reserve respond to the state of the economy during the current policy tightening cycle? (ii) How does the Federal Reserve's performance during the tightening cycle compare with more dovish or hawkish monetary policy stances? (iii) Was the Federal Reserve behind the curve as suggested by Summers (2021b)? And, if so, at what cost?

6.1 Data and Model Specification

We use an RC-SVAR specified at a quarterly frequency that includes output growth (as measured by the log difference of real GDP), core inflation (as measured by the log difference of the price index of personal consumption expenditures excluding food and energy), the federal funds rate, the growth in the stock of money (as measured by the log difference of M2), and Moody's seasoned Baa corporate bond yield relative to the yield on 10-year Treasury

constant maturity.[17] Often, we will refer to the latter as the credit spread. We express all the variables in terms of quarterly percent changes. The sample runs from 1959:Q1 until 2023:Q2. As typically done when working with time-varying SVARs at a quarterly frequency, we include a constant and two lags. Thus, we have $n = 5$, $p = 2$, and $m = 11$ in this model.

Our prior over $\boldsymbol{\phi}^{RC}$ is described in Appendix B. For ease of exposition, we present a summary. We set $\mathbf{m}_{\boldsymbol{\beta}_1}$ equal to the maximum likelihood estimate of a time-invariant reduced-form VAR with the same variables, constant, and lags based on the first $T_0 = 40$ observations available in our sample. We denote such an estimate by $\hat{\mathbf{B}}$. We set $\mathbf{V}_{\boldsymbol{\beta}_1}$ equal to the unbiased estimator for the variance of $\hat{\mathbf{B}}$. To set the values for $\boldsymbol{m}_{\boldsymbol{\delta}_1}$, $\boldsymbol{m}_{\boldsymbol{\gamma}_1}$, $\boldsymbol{V}_{\boldsymbol{\delta}_1}$, and $\boldsymbol{V}_{\boldsymbol{\gamma}_1}$, first we let $\hat{\boldsymbol{\Sigma}}$ denote the maximum likelihood estimate of the variance of the residuals. Second we use the Delta method to set the values for $\boldsymbol{m}_{\boldsymbol{\delta}_1}$, $\boldsymbol{V}_{\boldsymbol{\delta}_1}$, $\boldsymbol{m}_{\boldsymbol{\gamma}_1}$, and $\boldsymbol{V}_{\boldsymbol{\gamma}_1}$. For the parameters governing the step sizes of the processes for $\boldsymbol{\beta}_t$, $\boldsymbol{\delta}_t$, and $\boldsymbol{\gamma}$ we set $\mathbf{V}_{\boldsymbol{\beta}} \sim \text{IW}(\bar{k}^2_{\mathbf{V}_{\boldsymbol{\beta}}} \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}} \bar{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}}, \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}})$, $V_{\delta,i} \sim \text{IG}\left(\frac{\bar{\nu}_{\delta,i}}{2}, \frac{k^2_{w,i}\bar{\nu}_{\delta,i}\Psi_{V_{\delta,i}}}{2}\right)$ for $i = 1, \ldots, n$, and $V_{\gamma,i} \sim \text{IG}\left(\frac{\bar{\nu}_{\gamma,i}}{2}, \frac{k^2_{v,i}\bar{\nu}_{\gamma,i}\Psi_{V_{\gamma,i}}}{2}\right)$ for $i = 1, \ldots, n_\gamma$ with: $\bar{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}} = \boldsymbol{V}_{\boldsymbol{\beta}_1}$, $\bar{k}_{\mathbf{V}_{\boldsymbol{\beta}}} = 0.01$, $\bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}} = nm + 2$, $\bar{\nu}_{\delta,i} = 3$, $\Psi_{V_{\delta,i}} = 4 \times e'_{i,n} \boldsymbol{V}_{\boldsymbol{\delta}_1} e_{i,n}$, $k_{w,i} = 0.1$, for $i = 1 \ldots, n$, and $\bar{\nu}_{\gamma,i} = 6$, $\Psi_{V_{\gamma,i}} = 4 \times e'_{i,n_\gamma} \boldsymbol{V}_{\boldsymbol{\gamma}_1} e_{i,n_\gamma}$, $k_{v,i} = 0.1$, for $i = 1 \ldots, n_\gamma$, where $e_{i,x}$ denotes the i-th column of an $x \times x$ identity matrix.

Our results are based on four independent MCMC chains obtained using Algorithm 3. Each chain consists of 20,000 draws. For each chain, we keep every tenth draw of the structural parameters after discarding the first 2,500 draws. We use 1,800 particles for the reduced-form parameters and 3,600 particles for the orthogonal matrices to approximate the posterior distribution of $\boldsymbol{A}_t$. Drawing from $(\boldsymbol{F}_t)_{t=1}^T$ given $(\boldsymbol{A}_t)_{t=1}^T$ is straightforward given that the identification schemes do not restrict $(\boldsymbol{F}_t)_{t=1}^T$.

6.2 Identification

We identify the parameters of one of the equations of the SVAR, which we label the monetary policy equation, incorporating the insights of the identification strategies proposed by Uhlig (2005) and Arias, Caldara and Rubio-Ramírez (2019), but we allow for time variation in the systematic component of monetary policy. This is critical because the Federal Reserve's

[17] Appendix D provides a detailed description of the data sources.

reaction function to economic conditions has undoubtedly changed over time. For example, in October 1979, the Fed abandoned the federal funds rate as its main policy instrument and adopted non-borrowed reserves targeting to fight inflation. Similarly, the federal funds rate took a secondary role in December 2008 when it hit the effective zero lower bound (ZLB), and quantitative easing (QE) took center stage. In addition, changes in Federal Reserve chairs and the composition of the Federal Open Market Committee (FOMC) may have led to changes in the reaction function of monetary policy over time, (see, e.g., Boivin, 2006; Coibion and Gorodnichenko, 2011).

To address these changes, we consider an identification scheme that disciplines the systematic component of monetary policy during periods in which the federal funds rate can be deemed to have been the primary policy tool and that remains agnostic during periods in which the Fed targeted non-borrowed reserves or was constrained by the effective ZLB. When the federal funds rate is not the main policy instrument, the identifying restrictions follow Uhlig (2005) and concentrate on a minimal number of impulse responses. When the federal funds rate is the main policy instrument, we maintain the restrictions on impulse responses and, in addition, we impose restrictions on the contemporaneous structural parameters of the monetary policy equation following Arias, Caldara and Rubio-Ramírez (2019).[18] Importantly, the federal funds rate is not the main policy instrument during the whole sample; hence, the restrictions on the contemporaneous structural parameters of the monetary policy equation cannot be imposed for all the periods.

More specifically, we assume that the first equation of the SVAR is the monetary policy equation and, abstracting from the constant and lagged variables, we write it as

$$\mathrm{r}_t = \underbrace{\psi_{\Delta y,t}\Delta y_t + \psi_{\pi,t}\pi_t + \psi_{\Delta m,t}\Delta m_t + \psi_{\mathrm{cs},t}\mathrm{cs}_t}_{\text{Systematic Component}} + \underbrace{\sigma_{r,t}\varepsilon_{r,t}}_{\text{Shock}}, \tag{14}$$

where r_t is the federal funds rate, Δy_t is output growth, π_t is inflation, Δm_t is the growth rate of money, $\varepsilon_{r,t}$ is the monetary policy shock, $\psi_{\Delta y,t} = -a_{t,21}a_{t,31}^{-1}$, $\psi_{\pi,t} = -a_{t,11}a_{t,31}^{-1}$, $\psi_{i,t} = -a_{t,31}a_{t,31}^{-1}$, $\psi_{\Delta m,t} = -a_{t,41}a_{t,31}^{-1}$, and $\sigma_{r,t} = a_{t,31}^{-1}$, with $a_{t,ij}$ denoting the (i,j) entry of $\mathbf{A}_t$. Sometimes we will

[18] Wolf (2020) shows that restricting the systematic component of monetary policy can avoid shock-masquerading issues inherent to set identification with too few sign restrictions.

refer to the coefficients $(\psi_{\Delta y,t}, \psi_{\pi,t}, \psi_{\Delta m,t}, \psi_{\text{cs},t})$ as contemporaneous elasticities. Importantly, this equation clarifies that the monetary policy shock represents a deviation from a policy rule. Let us introduce Restrictions 1 and 2:

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

Restriction 1, motivated by Uhlig (2005), is imposed for the entire sample. Uhlig's (2005) celebrated identification scheme consists of imposing minimal sign restrictions on impulse responses motivated by economic theory without restricting the question of interest. Accordingly, when using his approach to study the economic effects of monetary policy shocks on output, he assumed that a monetary policy shock that increases the federal funds rate does not cause an increase in prices or non-borrowed reserves. In line with his identification scheme, we assume that the contemporaneous impulse response of inflation and the growth rate of the stock of money is negative in response to a monetary policy shock that increases the federal funds rate upon impact.[19]

Restriction 2. *Consider the following restrictions on the coefficients of Equation* (14)*,* $\psi_{\Delta y,t} \in (0,4)$*,* $\psi_{\pi,t} \in (0,4)$*,* $\psi_{\Delta m,t} \in (0,4)$*, and* $\psi_{cs,t} \in (-4,0)$*.*

The cost of using such a minimal number of restrictions is that some impulse responses consistent with Restriction 1 could be associated with an implausible systematic component of monetary policy or be subject to the shock masquerading issue highlighted by Wolf (2020).[20] Restriction 2, inspired by Arias, Caldara and Rubio-Ramírez (2019), addresses these concerns. The signs and bounds on the contemporaneous reaction of the federal funds rate to output growth and inflation follow directly from Arias, Caldara and Rubio-Ramírez (2019). While the signs of these responses can be viewed as uncontroversial, the upper limit of the bounds is somewhat arbitrary. We set it to 4 to strike a balance between using conservative bounds and ruling out implausible monetary policy behavior. Turning to the restriction on the response

[19]The length of the restrictions could be extended for one quarter to match the restricted horizons in Uhlig (2005). We impose only one period to keep the restrictions to the smallest number possible.

[20]See also Kilian and Murphy (2012), who highlighted a related issue in an SVAR of the global market for crude oil.

of the federal funds rate to the growth rate of money, we impose a positive sign following the work of Leeper and Zha (2003). The upper bound is large enough to encompass the point estimate of the elasticity of the federal funds rate to money in their model. Finally, the restriction on the response of the federal funds rate to the corporate credit spread is inspired by the work of Caldara and Herbst (2019), who highlighted that this response is crucial to address misspecification concerns in the monetary policy equation. In line with their estimates, we restrict this response to negative and impose a lower bound to rule out implausible large negative responses. This bound is such that the range of possible values for $\psi_{\text{cs},t}$ includes the 90 percent credible sets for this coefficient reported by Caldara and Herbst (2019). Restriction 2 is imposed on all the periods in our sample except for 1979Q4:1982Q4, 2009Q1:2015Q3, and 2020Q2:2021Q4. The first period corresponds to changes in monetary policy announced by Chair Volcker in late 1979. The second and third periods correspond to the QE policies around the Great Recession and COVID-19. The rationale for not imposing Restriction 2 during these three periods is as follows. In the first of these periods, the Federal Reserve explicitly targeted non-borrowed reserves. Lindsey et al. (2013) provide an extensive analysis of the New Operating Procedures announced by Chair Volcker on October 6, 1979. Their analysis highlights that the credibility of the prevailing discount rate framework came under severe stress following a near-split decision on the discount-rate vote and that the FOMC considered that targeting non-borrowed reserves would provide it the necessary flexibility to control inflation by allowing significant changes in interest rates. By October 1982, with inflation in a sustained downward trajectory, the Fed abandoned non-borrowed reserves as the main policy instrument. The abandonment was communicated less clearly than its adoption, and the Fed was vague about the details of its operating procedures for most of the 1980s and early 1990s (see, e.g., Lindsey, 2003). Chappell Jr, McGregor and Vermilyea (2005) describe the post-1982 conduct of monetary policy as indirect targeting of the federal funds rate, which gradually moved to direct targeting of the federal funds rate as made clear by the FOMC statement of March 1997.

The post-1982 period of interest rate targeting was interrupted when the Fed lowered interest rates to the 0 to 0.25 percent range, hitting its effective ZLB, and embarked on a period of QE policies to address the negative economic consequences of the Great Recession of

2007 to 2009. The conduct of monetary policy consistent with Equation (14) reemerged only after the federal funds rate lift-off announced in December 2015. Similarly, the COVID-19 pandemic induced a new period of QE policies and kept the federal funds rate at its effective ZLB from 2020Q2 until 2021Q4.

6.3 Systematic Component and Monetary Policy Shock

In this section, we show the systematic component of monetary policy and the monetary policy shock. Let us begin with the former. Figure 1 shows the contemporaneous elasticities of the federal funds rate to output growth, inflation, money growth, and the credit spread, from 1969Q4 until 2023Q2 except for those periods in which Restriction 2 is not imposed. The solid blue lines depict the point-wise posterior medians, and the solid gray lines represent the 68 percent point-wise posterior probability bands. As seen, the contemporaneous elasticity of the funds' rate to output growth exhibits three peaks. The first occurred in 1974 during the chairmanship of Arthur Burns, and it captures the sharp decline in the federal funds rate in response to the 1974-1975 recession. The second peak occurred in 1983-1984 during the chairmanship of Paul Volcker. This may come as a surprise, as his tenure at the Fed is commonly viewed as squarely focused on combating inflation. Even so, during this period, the annualized real growth rate of GDP averaged 6.8 percent. To the extent that the FOMC under Volcker's leadership viewed high growth as posing a risk to the progress they had made on the inflation front, it is natural to find that the federal funds rate was more sensitive to inflation during this period. The third peak occurred in 2001 during the chairmanship of Alan Greenspan when the Fed cut interest rates sharply in the face of the 2001 recession: In December 2000, the federal funds rate was 6.4 percent, and it ended 2001 at 1.8 percent. Outside of these peaks, the contemporaneous elasticities of the federal funds rate have been between 0.02 and 0.15, averaging about 0.1, implying that, other things constant, a one percentage point increase in annualized GDP growth would lead to a rise of 0.1 percent in the federal funds rate (annualized).

Turning to the contemporaneous elasticity of the federal funds rate to inflation, in line with the conventional wisdom, we find that the Fed reacted more aggressively to inflation in

the early 1980s than during the Great Moderation. However, we find the highest elasticities in the early 1970s under the first years of Burns's tenure. This suggests that through the lens of our model, the political pressure that President Nixon exerted on the Fed during the early 1970s (see, e.g., Drechsel, 2023) did not manifest in a lower response to inflation. Instead, as we will discuss below, such dovish pressure appears to be reflected in the reaction of the federal funds rate to the corporate credit spreads. When looking at the magnitude of the responses throughout the estimation sample, we find that while, on average, the annualized federal funds rate increased by 0.3 percentage points in response to a one percentage point increase in annualized core inflation, the range of responses is wide, being above 0.5 during the early 1970s and the early years of Volcker's chairmanship to as low as 0.15 during Janet Yellen's leadership.

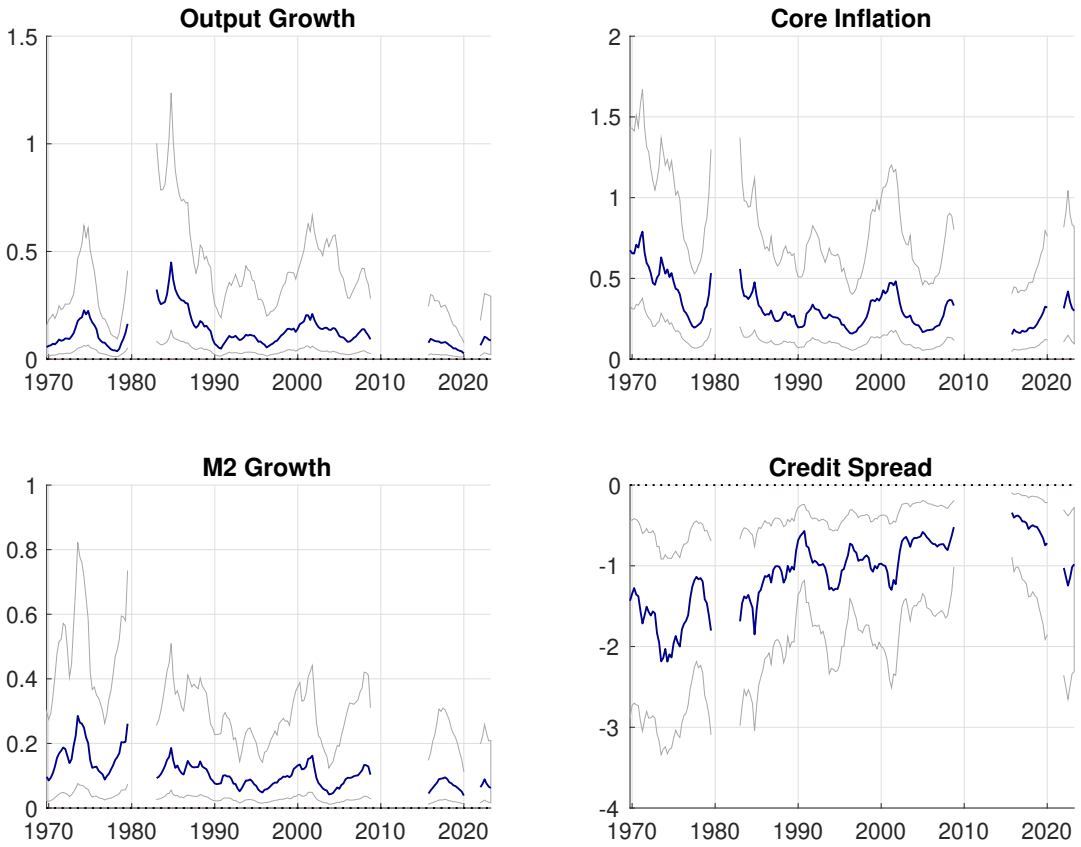


Figure 1: Contemporaneous Elasticities

Regarding the response of the federal funds rate to the growth rate of money, let us highlight that, on average, the elasticities are larger before 1979Q4 than after 1982Q4. This

is consistent with the view that policymakers in the early eighties were concerned with the reliability of monetary aggregates. For example, Lindsey et al. (2013) emphasize the following quote from Chairman Volcker obtained from an FOMC transcript from January 1980: "I would remind you that nothing that has happened—or that I've observed recently—makes the money/GNP relationship any clearer or more stable than before. Having gone through all these redefinition problems, one recognizes how arbitrary some of this is. It depends on how you define [money]".

Finally, we discuss the contemporaneous elasticity of the federal funds rate to corporate credit spreads. Analogously to Caldara and Herbst (2019), we find a significant reaction to changes in credit spreads. Nevertheless, we also find evidence of notable time variation in the magnitude of this coefficient. Interestingly, as highlighted above, the noticeable change in the contemporaneous elasticity of corporate credit spreads could be attributed to the political pressure faced by the Fed during the early 1970s. The coefficient moved from about −1.4 at the beginning of Burns's tenure to −2.2 at the end of Nixon's presidency. Hence, the tight credit spreads of the early 1970s induced the largest dovish pressure on the funds rate in our sample. Subsequently, it is clear that from the mid-1970s until the onset of the Great Recession, the Fed gradually became less responsive to corporate credit spreads. Perhaps not too surprisingly, the downward trend ended after the financial crisis. By mid-2023, the response was about -1, which aligns with the posterior median estimates for the 1990s and early 2000s. While Caldara and Herbst (2019) focus on the period 1994-2007, our results indicate that corporate credit spreads played a potentially even more critical role in the conduct of monetary policy during the 1970s and 1980s.

Figure 2 shows the standard deviation of the monetary policy shock. In line with other estimates in the literature, it has declined since the early 1980s. Notably, the clear evidence of time variation in the systematic component of monetary policy and the standard deviation of the monetary policy shock suggests that a time-varying structural model could be critical to studying the conduct of monetary policy in the U.S. over a long sample.

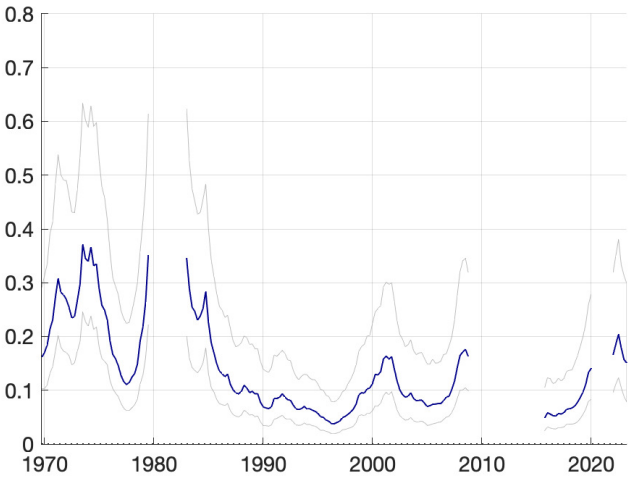


Figure 2: Standard Deviation of the Monetary Policy Shock

6.4 Dissecting the 2022-2023 Tightening Cycle

As outlined in the introduction, it is widely acknowledged that the Federal Reserve initiated a tightening of monetary policy in early 2022 due to inflation concerns (see Romer and Romer, 2023b). We now examine the degree to which the unexpected changes in the federal funds rate from the second quarter of 2022 to the second quarter of 2023 are attributable to either the systematic component of monetary policy or monetary policy shocks. Figure 3 presents our model-based forecasts for the period 2022Q2-2023Q2, using data from 2022Q1 (represented by a dotted line). Thus, forecast and shock contributions are constructed using point-wise means conditional on the distribution of structural parameters corresponding to 2022Q1. This figure also illustrates the cumulative impact of structural shocks on the unexpected variations in the federal funds rate, output growth, and core inflation for each quarter between 2022Q2 and 2023Q2. We depict monetary policy shocks using red bars, while the contributions from "Non-Monetary Policy Shocks" are shown in yellow bars. Both the forecast and shock contributions are calculated using point-wise means.

The forecast for the federal funds rate shows that the model significantly under-predicted its trajectory. Overall, the forecast for output growth was also lower than in the data. In contrast, the projection for inflation was remarkably accurate. We contrast our model with the median predictions from the 2022Q1 and 2022Q2 Survey of Professional Forecasters

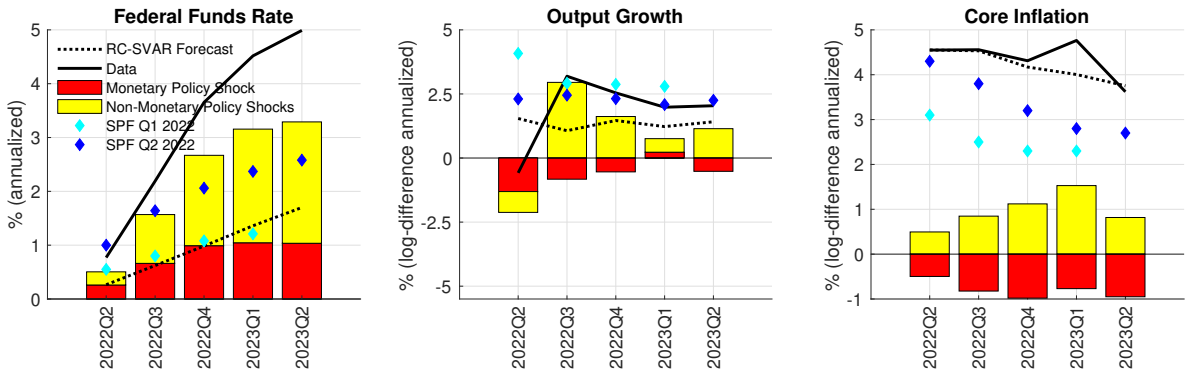


Figure 3: Non-Monetary versus Monetary Policy Shocks

(SPF).[21] The 2022Q1 SPF participants are at an informational disadvantage relative to our model, while the 2022Q2 SPF participants are at an informational advantage compared to our model. The main takeaway from this comparison is that the model forecasts are broadly in line with the SPF for the case of the federal funds rates, while the SPF does better than the model when forecasting output growth and worse in the case of inflation.

Looking at the decomposition of the federal funds rate, it is clear that most of the unexpected changes in the federal funds rate can be attributed to the systematic component of monetary policy. In particular, interest rates increased beyond what could have been predicted in 2022Q1, mainly as a systematic response to non-monetary policy shocks. This can be seen in the figure by comparing the height of the yellow bars to the height of the red bars. The yellow bars represent the contribution of the systematic part of monetary policy to unexpected changes in the federal funds rate. In contrast, the red bars represent the contribution of the monetary policy shocks. Consequently, the lion's share of the unpredictable increases in the interest rate can be attributed to the Fed's policy reaction function. This is in line with the findings of the literature when analyzing other periods, which argues that most of the variation in policy instruments is due to the systematic component of policy and not monetary policy shocks (for example, see Sims, 1998). Even so, monetary policy shocks also played a role, amounting to about 100 basis points of the unexpected change in the federal funds rate by 2023Q2. Table 1 shows the details of this decomposition for the case of the fed funds rate.

As shown in the figure, non-monetary policy shocks are the main contributor to the

[21] We interpret the SPF projections for the 3-month T-bill rate as projections for the federal funds rate.

	2022Q2	2022Q3	2022Q4	2023Q1	2023Q2
Predictable	0.27	0.62	0.98	1.36	1.70
Unpredictable due to Systematic	0.24	0.91	1.68	2.11	2.25
Unpredictable due to MP Shocks	0.26	0.66	0.99	1.04	1.04
Federal Funds Rate	0.77	2.19	3.65	4.52	4.99

Table 1: Federal Funds Rate Decomposition (p.p.)

unexpected changes in output growth and inflation: non-monetary policy shocks caused the economy to run hotter (larger than expected output growth and inflation) than predicted. This explains why the systematic part of monetary policy contributed positively to the unexpected change in the fed fund rate.

6.5 Romer and Romer's July 2022 Monetary Policy Shock

To cross-verify our results, we now analyze if the monetary shocks that we find between 2022Q2 and 2023Q2 coincide with those found with Romer and Romer (2023a).[22] In particular, their paper argues that there was a contractionary monetary policy shock in July 2022. Our estimates are consistent with such a narrative. This can be seen in Figure 4, which shows the series of monetary shocks identified by our model together with Romer and Romer's (2023a) shock between 2022Q1 and 2023Q2. The black line shows the point-wise posterior medians for the identified monetary policy shock, while the green areas represent the 68 percent point-wise posterior probability bands. According to our estimates, there is a contractionary shock in 2022Q3.

Even so, our estimates do not perfectly match the narrative record of the latest tightening cycle in Romer and Romer (2023a). Whereas they saw some signs but not definite evidence of a contractionary monetary policy shock in 2022Q2, our model detects a contractionary shock in such a quarter. Looking beyond 2022, we cannot compare our approach with theirs because their analysis was conducted in January 2023. Nevertheless, our model estimates show some evidence of an expansionary monetary policy shock in 2023Q1, during which the

[22]We think that this comparison could be helpful for some readers even though, as highlighted by Sims and Zha (2006a), the resulting shock series are not directly comparable because of differences in the treatment of the systematic component of monetary policy.

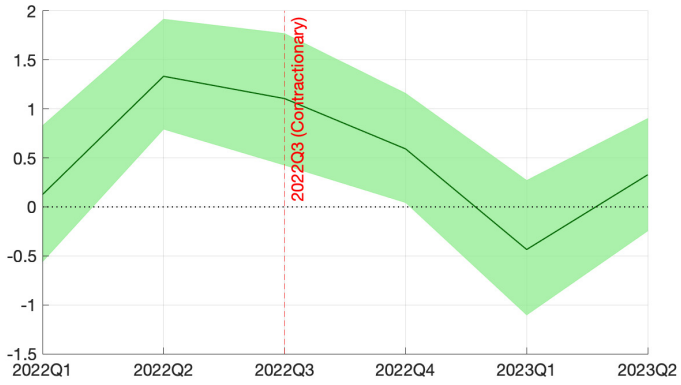


Figure 4: Monetary Policy Shocks

Fed decreased the pace of interest rate increases from 50 to 25 basis points.

6.6 Policy Counterfactual Simulations

One of the benefits of considering a structural model is the ability to look at the effects of counterfactual experiments that change the structural parameters in the monetary policy equation. For instance, Sims and Zha (2006b) look at impulse responses to shocks in an SVAR where the estimated policy equation is replaced by one in which the monetary authority is unresponsive to other variables in the system. Similarly, Primiceri (2005) conducts an experiment he calls "planting Greenspan into the 1970s." The idea of such an experiment is to replay history, drawing the parameters of the policy rule in the 1970s from their posterior in 1991-1992, to assess the consequences of a change in the systematic policy component. Building on this tradition, we produce two types of counterfactuals. In the first type, we modify the systematic component of monetary policy, keeping other aspects of the model unchanged. This approach aligns with the framework presented in Sims and Zha (2006a), wherein rational agents cannot comprehend or anticipate policy changes. In the second type, we generate counterfactuals inspired by the work of McKay and Wolf (2023) and Caravello, McKay and Wolf (2023). These counterfactuals address the expectational concerns of the Lucas critique.

Let us begin with the counterfactuals relying on a policy change to the systematic component of monetary policy. Specifically, we conduct two counterfactual simulations that

replay history since 2022Q2, assuming that the FOMC would have responded to contemporary inflation differently than what would be prescribed by our estimated policy rule. In the first simulation, which we label Hawkish Fed, we replace the model's estimated reaction to contemporaneous inflation with a twice as large response. In the second simulation, which we label Dovish Fed, we replace the model's estimated reaction to contemporaneous inflation with a response that is half as large.

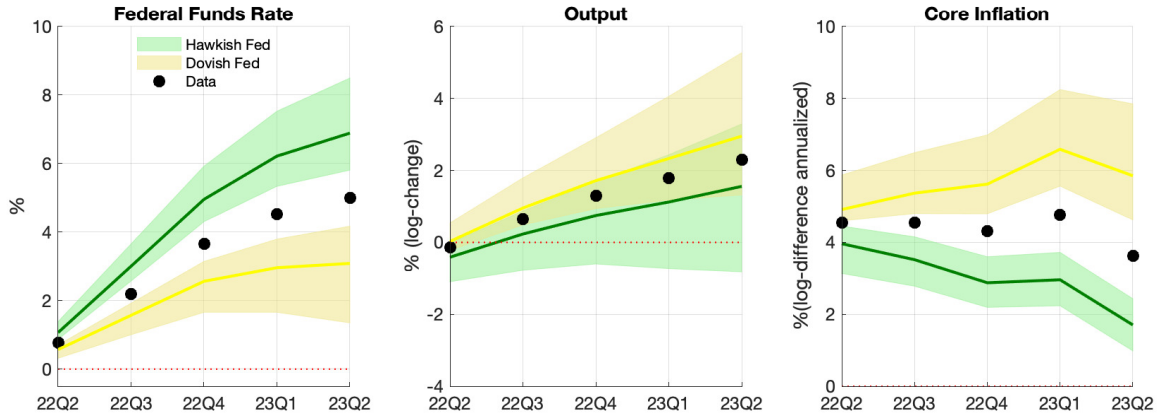


Figure 5: Counterfactuals

Figure 5 presents the results. Focusing on posterior medians, through the lens of our model, under the Dovish Fed counterfactual, the economy would have marginally overheated with output increasing above the pace implied by the U.S. Congressional Budget Office's estimates of potential GDP, and inflation would have run persistently above 5 percent. Under the Hawkish Fed counterfactual, inflation would have quickly decreased in terms of economic activity at a small cost. In the Hawkish counterfactual, the output in the second quarter of 2023 would have been about 0.7 percent lower. When looking at the output level at risk, the lower envelope of the 68 percent probability bands shows that the cost in terms of output could have been as high as 3.1 percent.

While the Bayesian paradigm adopted by our approach mitigates some of the concerns associated with the expectational issues of the Lucas critique, it is still possible that economic agents would become aware of the policy change over time, potentially affecting the lessons obtained from the counterfactuals. To address the latter, we conduct a variant of the Hawkish Fed and Dovish Fed counterfactuals that rely neither on changing the systematic component

of monetary policy nor on ex-post monetary policy surprises. For this reason, we refer to these counterfactuals as robust to expectational concerns. Specifically, we construct a Hawkish Fed (Dovish Fed) counterfactual by replaying history since 2022Q2, assuming that on such quarter, there was an additional contemporaneous monetary policy shock that other things constant would have led to a median increase (decrease) of 75 basis point in the federal funds rate. We choose 75 basis points so the median federal funds rate does not exceed zero in the Dovish Fed counterfactual. Figure 6 shows the results, which, as can be seen, are broadly similar to those obtained under the first set of counterfactuals.

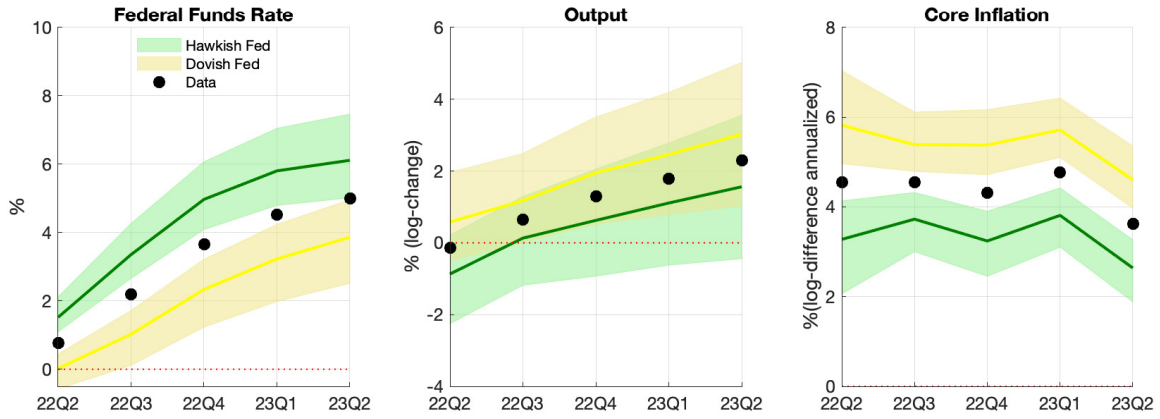


Figure 6: Counterfactuals Robust to Expectational Concerns

6.7 Was the Fed Behind the Curve?

Thus far, our analysis has revolved around understanding the 2022-2023 tightening cycle since its onset. This section focuses on whether the Federal Reserve was late to increase interest rates. This is an interesting question because, in early 2021, influential economists expressed concerns that the American Rescue Plan Act (signed into law on March 11, 2021) could result in a surge of inflation not seen since the 1970s unless the Federal Reserve responded to the program (see, e.g. Summers, 2021a; Blanchard, 2021).

Figure 7 shows our model-based forecasts for 2021Q2:2021Q4 based on 2021Q1 data (dotted line). In addition, the figure shows the cumulative contribution of structural shocks to the unexpected change in the federal funds rate for each quarter from 2021Q2 until 2021Q4. As above, we plot monetary policy shocks (red bars), and the rest of the structural shocks

are aggregated into a "Non-Monetary Policy Shocks" category (yellow bars). Forecast and shock contributions are constructed using point-wise means conditional on the distribution of structural parameters corresponding to 2021Q1.

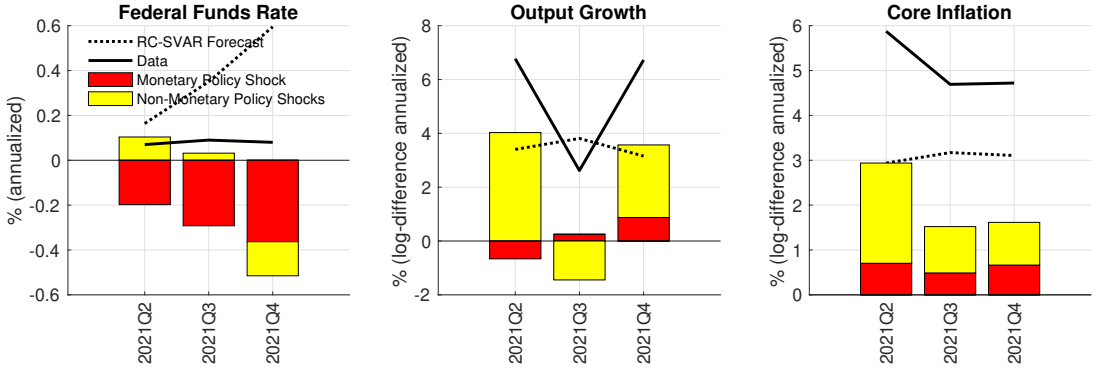
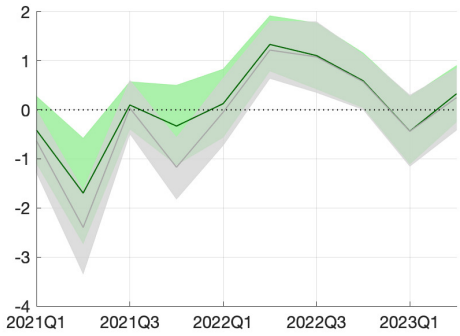


Figure 7: Monetary Policy versus Non-Monetary Policy Shocks

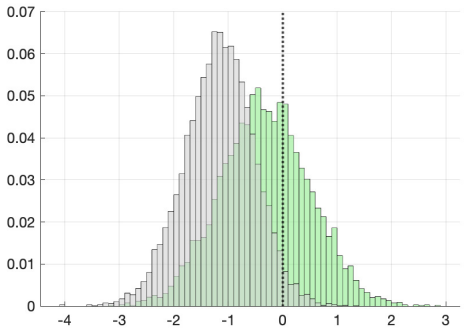
Let us begin by discussing the point-wise mean forecasts. The RC-SVAR predicted an earlier-than-realized lift-off for the federal funds rate: the model expected the funds rate to be about 0.6 percent by the end of 2021. This projection was associated with a prediction for output growth of nearly 4 percent in 2021Q1, slowing to slightly above 3 percent by 2021Q4, and a prediction for inflation of about 3 percent throughout the forecast horizon. Regarding the shock decomposition, the cumulative unexpected change in the federal funds rate is driven by expansionary monetary policy shocks supporting the view that the monetary policy stance was accommodative, and the FMOC fell behind the curve during this period. The inflationary consequences of these actions can be seen in the cumulative decomposition of the unexpected change in core inflation: monetary policy shocks on average contributed about 0.6 percentage points to annualized core inflation during the period under analysis. The remaining and larger share of the unexpected change in inflation can be attributed to non-monetary policy shocks. Thus, interestingly, we find that although the Fed was running behind the curve, this was not the primary factor underlying the inflation run-up. The relative contribution of monetary policy shocks to the unexpected change in the output growth rate is small.

6.8 Do Time-Varying Sign Restrictions Matter?

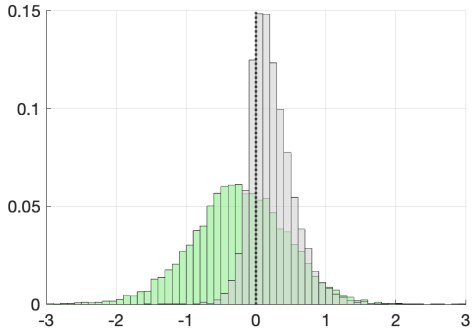
We conclude our empirical application by demonstrating that the ability to impose sign restrictions in only selected periods can affect inference. To this end, we contrast the monetary policy shocks obtained under the identification scheme in Section 6.2 (which in this section we will refer to as the baseline identification scheme) with an alternative identification scheme in which Restrictions 1 and 2 are imposed in all periods of our sample.[23]



(a) Monetary Policy Shocks



(b) Histogram Monetary Policy Shocks



(c) Impact Response of Output

Figure 8: Monetary Policy Shocks

Note: Panel (a): Point-wise posterior medians and the point-wise 68 percent posterior probability bands for the estimated monetary policy shocks between 2021Q1 and 2023Q2 when using the baseline identification scheme (solid dark green lines for the median and light green shades for the bands) and when using the alternative identification scheme (solid dark gray lines for the median and light gray shades for the bands). Panels (b) and (c): The green (gray) histogram corresponds to the baseline (alternative) identification scheme.

Panel (a) in Figure 8 shows the point-wise posterior medians and the point-wise 68 percent

[23]The results of the alternative identification scheme are based on 10,000 draws from the posterior distribution obtained using Algorithm 3.

posterior probability bands for the estimated monetary policy shocks between 2021Q1 and 2023Q2 when using the baseline identification scheme (solid dark green lines for the median and light green shades for the bands) and when using the alternative identification scheme (solid dark gray lines for the median and light gray shades for the bands). We focus on the period 2021Q1:2023Q2 because during this time, clear voices were calling for monetary policy action (see, e.g. Blanchard, 2021; Summers, 2021a,b), and it is also the period in which we find the most significant discrepancies between the two identification schemes under analysis.

Although the implied monetary policy shocks over the period under analysis are similar under both identification schemes, there are some differences. Through the lens of the alternative identification scheme, there was an expansionary monetary policy shock in 2021Q4, while the baseline finds that the monetary policy shock was nearly centered around zero. The difference can be seen in Panel (b), where we plot the histograms of the estimated monetary policy shocks in 2021Q4 for the baseline (green histogram) and the alternative (gray histogram) identification schemes. In addition, the implications for output also differ depending on the identification scheme in place. Panel (c) shows the histograms of the posterior estimates of the contemporaneous impulse response of real GDP to a one-standard-deviation expansionary monetary policy shock in 2021Q2. As can be seen, using the alternative identification scheme, the expansionary monetary policy shock of 2021Q2 is more likely to have had positive effects on output. This highlights that differences in impulse responses could emerge, even when both schemes find similar monetary policy shocks, as in 2021Q2.

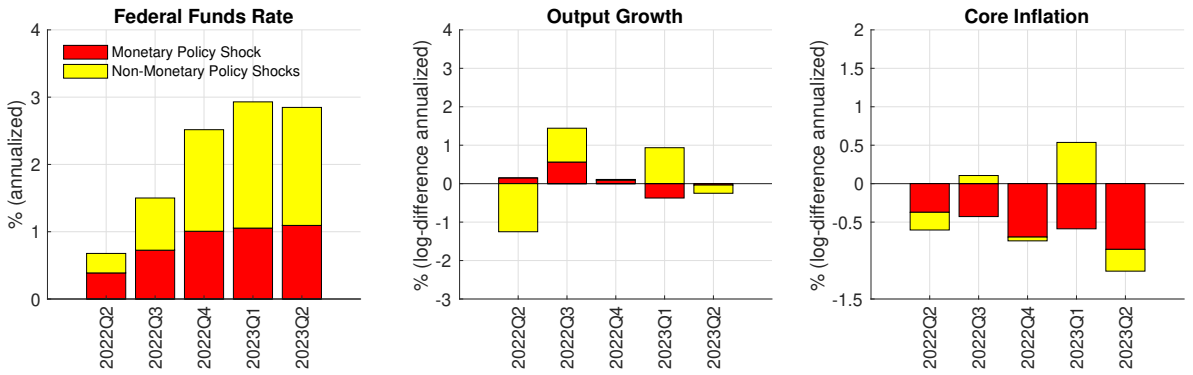
Overall, the results presented in Figure 8 illustrate that the ability to impose the sign restrictions on selected periods can affect posterior inference about the effects of monetary policy at a critical juncture. More broadly, time-varying sign restrictions offer a helpful toolkit for those interested in adhering on a time-specific basis to the *If you know it, impose it! If you do not know it, do not impose it!* principles for inference based on sign restrictions outlined in Uhlig (2017).

6.9 Constant Parameters SVAR

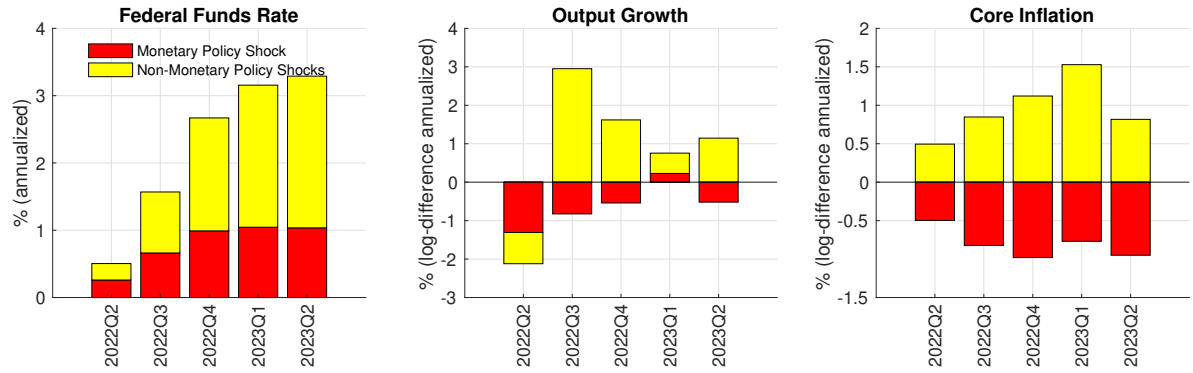
Researchers employing a constant parameters SVAR to conduct monetary policy analysis in our sample, while adhering to the aforementioned identification principles, would be hesitant to impose Restriction 2. This reluctance stems from the fact that the federal funds rate was not the primary policy instrument throughout the entire period under analysis. Consequently, these researchers would be confident only in imposing Restriction 1, as typically done following the work of Uhlig (2005). In this section, we explore the consequences of the inability to selectively impose Restriction 2 by contrasting the historical decomposition presented in Section 6.4 with the one obtained using a constant parameters SVAR identified with Restriction 1. To this end, we specify an SVAR at a quarterly frequency featuring the same variables and sample as in the RC-SVAR. We include a constant and four lags, as typically done in this setting when working with constant parameters, and we use the "weak" priors described in Uhlig (2005).[24] Figure 9 presents the comparison. Panel (a) shows the cumulative impact of structural shocks on the unexpected variations in the federal funds rate, output growth, and core inflation for each quarter between 2022Q2 and 2023Q2 for the constant parameters case. Panel (b) reproduces the results for the RC-SVAR shown in Figure 3. As we will discuss below, there are two salient differences between the historical decompositions implied by the constant parameters SVAR and the RC-SVAR, respectively.

First, through the lens of the former, monetary policy shocks contribute insignificantly to the unexpected variations in output growth relative to the RC-SVAR. Therefore, a constant parameters SVAR, identified solely by Restriction 1, presents a markedly different picture regarding the drivers of unexpected output growth fluctuations between 2022Q2 and 2023Q2. Whereas the RC-SVAR indicates that monetary policy shocks decelerated output growth in the face of non-monetary policy shocks that caused the economy to run hotter, the constant parameters variant suggests that monetary policy shocks propelled output growth in 2022 and were roughly neutral after that. Figure 10 sheds light on this discrepancy by comparing the impulse responses to a monetary policy shock in the RC-SVAR with those obtained in the constant parameters SVAR. The solid dark green lines and the light green shades

[24] The results of the constant parameters SVAR are based on 10,000 draws from the posterior distribution obtained using Algorithm 1 in Arias, Rubio-Ramírez and Waggoner (2018).



(a) Constant Parameters SVAR



(b) RC-SVAR

Figure 9: Monetary Policy vs Non-Monetary Policy Shocks

correspond to the point-wise posterior median and 68 percent point-wise probability bands for the RC-SVAR. The solid and dotted black lines correspond to the point-wise posterior median and 68 percent point-wise probability bands for the constant parameters SVAR. In the case of the RC-SVAR, we focus on 2022Q3—the quarter associated with the contractionary shock identified by Romer and Romer (2023a)—and we compute the impulse responses to a unit standard deviation monetary policy shock as in Primiceri (2005). In the case of the constant parameters SVAR, we scale the shock such that upon impact, the median increase in the federal funds rate equals the median increase in the RC-SVAR. When using the RC-SVAR, the posterior median of output is negative following a contractionary monetary policy shock, and, as a result, the positive monetary policy shocks needed to explain the unexpected increase in the fed funds rate negatively affect output. In contrast, in the constant parameters SVAR, the posterior median response of output is positive, indicating a contractionary monetary

policy shock. The positive posterior median response of output is in line with the results in Uhlig (2005).

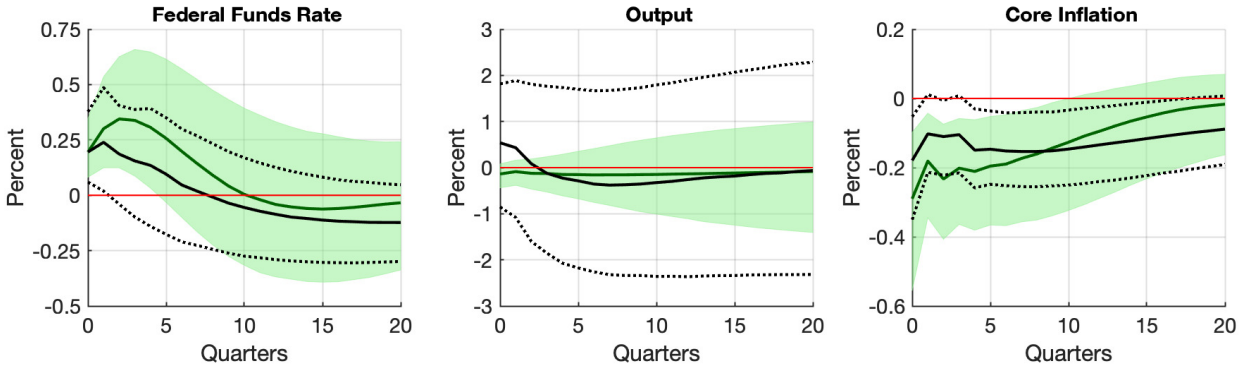


Figure 10: Selected Impulse Responses to Monetary Policy Shocks

Note: The solid dark green lines and the light green shades correspond to the point-wise posterior median and 68 percent point-wise probability bands for the RC-SVAR. The solid and dotted black lines correspond to the point-wise posterior median and 68 percent point-wise probability bands for the constant parameters SVAR.

Second, although the non-monetary policy shocks explain the largest share of the unexpected variation of the federal funds rate in both models, in the constant parameters SVAR, these shocks do not operate through output and inflation. This can be seen by comparing Panels (a) and (b) in Figure 9: the yellow bars are much less important in explaining the unexpected variation of output growth and inflation in the constant parameters case than in the RC-SVAR. Thus, researchers employing constant parameters SVAR would obtain little support for the view that unexpectedly higher interest rates were driven by non-monetary policy shocks that induced the economy to run hotter. Instead, as shown in Figure 11 depicting the contributors to unexpected variations of money growth and credit spreads for the constant parameters SVAR, these researchers would conclude that non-policy shocks induced unexpected variation in both the federal funds rate and money growth without a clear association with the real or financial side of the economy.

In sum, these two differences show how the ability to impose time-specific sign restrictions can shape researchers' views on classical macroeconomic questions, such as what monetary policy does.

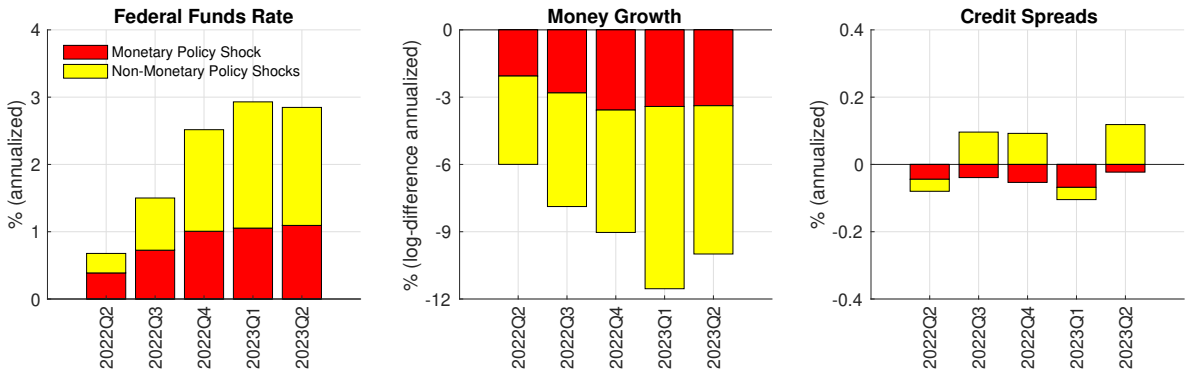


Figure 11: Non-Monetary versus Monetary Policy Shocks

7 Conclusion

The theory developed in this paper can be extended in multiple directions. First, it offers a path forward to researchers interested in conducting empirical work using time-varying SVAR with priors that assign the same density to observationally equivalent sequences of structural parameters. Second, the techniques can be adapted to consider zero restrictions, provided that one considers the volume elements when inducing priors over sequences of structural parameters. This extension is not straightforward since these volume elements are more intricate than the constant parameters case considered by Arias, Rubio-Ramírez and Waggoner (2018). Third, exploring the role of inference about the fixed parameters is another interesting line of research. Finally, given that our methodology is compatible with a wide range of time-varying models, it could be fruitful to conduct marginal likelihood comparisons across them conditional on the identifying restrictions.

References

Aastveit, K. A., F. Furlanetto, and F. Loria (2023). Has the Fed Responded to House and Stock Prices? A Time-Varying Analysis. *Review of Economics and Statistics 105*(5), 1314–1324.

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang (2016). Drifts and Volatilities Under

Measurement Error: Assessing Monetary Policy Shocks over the Last Century. *Quantitative Economics 7*(2), 591–611.

Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle Markov Chain Monte Carlo Methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72*(3), 269–342.

Archakov, I. and P. R. Hansen (2021). A New Parametrization of Correlation Matrices. *Econometrica 89*(4), 1699–1715.

Arias, J. E., D. Caldara, and J. F. Rubio-Ramírez (2019). The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure. *Journal of Monetary Economics 101*, 1–13.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2018). Inference Based on Structural Vector Autoregressions Identified with Sign and Zero Restrictions: Theory and Applications. *Econometrica 86*(2), 685–720.

Arias, J. E., J. F. Rubio-Ramírez, and M. Shin (2023). Macroeconomic Forecasting and Variable Ordering in Multivariate Stochastic Volatility Models. *Journal of Econometrics 235*(2), 1054–1086.

Asai, M. and M. McAleer (2009). The Structure of Dynamic Correlations in Multivariate Stochastic Volatility Models. *Journal of Econometrics 150*(2), 182–192.

Baumeister, C. and G. Peersman (2013). Time-Varying Effects of Oil Supply Shocks on the US Economy. *American Economic Journal: Macroeconomics 5*(4), 1–28.

Blanchard, O. (2021). In Defense of Concerns over the 1.9 Trillion Relief Plan. *Peterson Institute for International Economics*.

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*(18-11), 1–61.

Boivin, J. (2006). Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data. *Journal of Money, Credit and Banking 38*(5), 1149–1173.

Brunnermeier, M., D. Palia, K. Sastry, and C. Sims (2021). Feedbacks: Financial Markets and Economic Activity. *American Economic Review 111*, 1845–1879.

Caldara, D. and E. Herbst (2019). Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs. *American Economic Journal: Macroeconomics 11*(1), 157–92.

Caravello, T. E., A. McKay, and C. K. Wolf (2023). Evaluating Policy Counterfactuals: A "VAR-Plus" Approach. *Working Paper*.

Carter, C. K. and R. Kohn (1994). On Gibbs Sampling for State Space Models. *Biometrika 81*(3), 541–553.

Chan, J. C., G. Koop, and X. Yu (2021). Large Order-Invariant Bayesian VARs with Stochastic Volatility. *arXiv preprint arXiv:2111.07225*.

Chappell Jr, H. W., R. R. McGregor, and T. A. Vermilyea (2005). *Committee Decisions on Monetary Policy: Evidence from Historical Records of the Federal Open Market Committee*. The MIT Press.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1996). The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds. *Review of Economics and Statistics*, 16–34.

Clarida, R., J. Gali, and M. Gertler (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *Quarterly Journal of Economics 115*(1), 147–180.

Coibion, O. and Y. Gorodnichenko (2011). Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation. *American Economic Review 101*(1), 341–70.

Debortoli, D., J. Galí, and L. Gambetti (2020). On the Empirical (Ir)Relevance of the Zero Lower Bound Constraint. *NBER Macroeconomics Annual 34*, 141–170.

Drechsel, T. (2023). Estimating the Effects of Political Pressure on the Fed: A Narrative Approach with New Data. *Working Paper*.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods*. Oxford University Press.

Galí, J. and L. Gambetti (2015). The Effects of Monetary Policy on Stock Market Bubbles: Some Evidence. *American Economic Journal: Macroeconomics 7*(1), 233–57.

Giannone, D., M. Lenza, and G. E. Primiceri (2015). Prior Selection for Vector Autoregressions. *Review of Economics and Statistics 97*(2), 436–451.

Herwartz, H. and H. Lütkepohl (2014). Structural Vector Autoregressions with Markov Switching: Combining Conventional with Statistical Identification of Shocks. *Journal of Econometrics 183*(1), 104–116.

Hubrich, K. and D. F. Waggoner (2022). The Transmission of Financial Shocks and Leverage of Financial Institutions: An Endogenous Regime Switching Framework.

Kilian, L. and D. P. Murphy (2012). Why Agnostic Sign Restrictions Are Not Enough: Understanding the Dynamics of Oil Market VAR Models. *Journal of the European Economic Association 10*(5), 1166–1188.

Lanne, M. and H. Lütkepohl (2008). Identifying Monetary Policy Shocks via Changes in Volatility. *Journal of Money, Credit and Banking 40*(6), 1131–1149.

Lanne, M., H. Lütkepohl, and K. Maciejowska (2010). Structural Vector Autoregressions with Markov Switching. *Journal of Economic Dynamics and Control 34*(2), 121–131.

Leeper, E. M. and T. Zha (2003). Modest Policy Interventions. *Journal of Monetary Economics 50*(8), 1673–1700.

Lindsey, D. E. (2003). A Modern History of FOMC Communication: 1975-2002. *Document authorized for public release by the FOMC Secretariat on 05/27/2010*.

Lindsey, D. E., A. Orphanides, R. H. Rasche, et al. (2013). The Reform of October 1979: How It Happened and Why. *Federal Reserve Bank of St. Louis Review 95*(6), 487–542.

Lindsten, F., M. I. Jordan, and T. B. Schon (2014). Particle Gibbs with Ancestor Sampling. *Journal of Machine Learning Research 15*, 2145–2184.

Lütkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag.

Lütkepohl, H. and A. Netšunajev (2017). Structural Vector Autoregressions with Heteroskedasticity: A Review of Different Volatility Models. *Econometrics and Statistics 1*, 2–18.

McKay, A. and C. K. Wolf (2023). What Can Time-Series Regressions Tell Us About Policy Counterfactuals? *Econometrica 91*(5), 1695–1725.

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*.

Prado, R. and M. West (2010). *Time Series: Modeling, Computation, and Inference*. CRC Press.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies 72*(3), 821–852.

Rigobon, R. (2003). Identification Through Heteroskedasticity. *Review of Economics and Statistics 85*(4), 777–792.

Rigobon, R. and B. Sack (2003). Measuring the Reaction of Monetary Policy to the Stock Market. *Quarterly Journal of Economics 118*(2), 639–669.

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

Romer, C. D. and D. H. Romer (2023b). Presidential Address: Does Monetary Policy Matter? The Narrative Approach after 35 Years. *American Economic Review 113*(6), 1395–1423.

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

Rubio-Ramírez, J., D. Waggoner, and T. Zha (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *Review of Economic Studies 77*(2), 665–696.

Sentana, E. and G. Fiorentini (2001). Identification, Estimation and Testing of Conditionally Heteroskedastic Factor Models. *Journal of Econometrics 102*(2), 143–164.

Sims, C. (1998). Comment on Glenn Rudebusch's "Do Measures of Monetary Policy in a VAR Make Sense?". *International Economic Review 39*(4), 933–41.

Sims, C. A. and T. Zha (2006a). Does Monetary Policy Generate Recessions? *Macroeconomic Dynamics 10*(2), 231–272.

Sims, C. A. and T. Zha (2006b). Were There Regime Switches in US Monetary Policy? *American Economic Review*, 54–81.

Summers, L. (2021a). The Biden Stimulus Is Admirably Ambitious. But It Brings Some Big Risks, Too. *The Washington Post*. February 4, 2021.

Summers, L. (2021b). On Inflation, It's Past Time for Team 'Transitory' to Stand Down. *The Washington Post*. November 15, 2021.

Uhlig, H. (1994). On Singular Wishart and Singular Multivariate Beta Distributions. *The Annals of Statistics 22*(1), 395–405.

Uhlig, H. (2005). What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure. *Journal of Monetary Economics 52*(2), 381–419.

Uhlig, H. (2017). Shocks, Sign Restrictions, and Identification. *Advances in Economics and Econometrics 2*, 95.

Wolf, C. K. (2020). SVAR (Mis-)Identification and the Real Effects on Monetary Policy. *American Economic Journal: Macroeconomics 12*(4), 1–32.

Appendix

A Proofs

Proof of Proposition 1. The likelihood is given by

$$p((\mathbf{y}_t)_{t=1}^T \mid \mathbf{x}_1, (\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}) = \prod_{t=1}^T p(\mathbf{y}_t \mid \mathbf{x}_t, \mathbf{A}_t, \mathbf{F}_t),$$

where $p(\mathbf{y}_t \mid \mathbf{x}_t, \mathbf{A}_t, \mathbf{F}_t)$ is Gaussian with mean $\mathbf{x}_t' \mathbf{F}_t \mathbf{A}_t^{-1}$ and variance $(\mathbf{A}_t \mathbf{A}_t')^{-1}$. Note here we are being explicit about the fact that we are conditioning on the initial conditions. So, if there exists $(\mathbf{Q}_t)_{t=1}^T \in \mathcal{O}_n^T$ such that $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T = (\mathbf{A}_t \mathbf{Q}_t, \mathbf{F}_t \mathbf{Q}_t)_{t=1}^T$, then the likelihoods at $((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi})$ and $((\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T, \tilde{\boldsymbol{\phi}})$ are equal for all $(\mathbf{y}_t)_{t=1}^T$. Thus, $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ and $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T$ are observationally equivalent.

Now assume that $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ and $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T$ are observationally equivalent. Again, because $p(\mathbf{y}_t \mid \mathbf{x}_t, \mathbf{A}_t, \mathbf{F}_t)$ is Gaussian with mean $\mathbf{x}_t' \mathbf{F}_t \mathbf{A}_t^{-1}$ and variance $(\mathbf{A}_t \mathbf{A}_t')^{-1}$ and $p(\mathbf{y}_t \mid \mathbf{x}_t, \tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)$ is Gaussian with mean $\mathbf{x}_t' \tilde{\mathbf{F}}_t \tilde{\mathbf{A}}_t^{-1}$ and variance $(\tilde{\mathbf{A}}_t \tilde{\mathbf{A}}_t')^{-1}$, it must be the case that $(\mathbf{A}_t \mathbf{A}_t')^{-1} = (\tilde{\mathbf{A}}_t \tilde{\mathbf{A}}_t')^{-1}$ and $\mathbf{x}_t' \mathbf{F}_t \mathbf{A}_t^{-1} = \mathbf{x}_t' \tilde{\mathbf{F}}_t \tilde{\mathbf{A}}_t^{-1}$. The former implies that $(\mathbf{A}_t^{-1} \tilde{\mathbf{A}}_t)(\mathbf{A}_t^{-1} \tilde{\mathbf{A}}_t)' = \boldsymbol{I}_n$, so that $\mathbf{Q}_t \equiv \mathbf{A}_t^{-1} \tilde{\mathbf{A}}_t \in \mathcal{O}_n$. The latter implies that it must be the case that $\mathbf{x}_t'(\mathbf{F}_t \mathbf{Q}_t - \tilde{\mathbf{F}}_t) = \mathbf{0}$ for almost every $\mathbf{x}_t$. For $t > p$, the span of the support of $\mathbf{x}_t$ is all of $\mathbb{R}^m$. However, for $1 \le t \le p$, the span of the support of $\mathbf{x}_t$ depends on the assumptions about $\mathbf{x}_1$. While $\mathbf{x}_1$ is observed, it is a random variable, and we assume that the span of the support of $\mathbf{x}_1$ is all of $\mathbb{R}^m$. This implies that the span of the support of $\mathbf{x}_t$ is all of $\mathbb{R}^m$ for $1 \le t \le T$. Thus, $\mathbf{F}_t \mathbf{Q}_t = \tilde{\mathbf{F}}_t$ for $1 \le t \le T$. □

Proof of Proposition 2. Because $v_{f_h^{-1}}((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T)$ does not depend on $(\mathbf{Q}_t)_{t=1}^T$, the conditional prior $p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi})$ satisfies Equation (3) if and only if the equivalent conditional prior $p_R((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi})$ satisfies

$$p_R((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi}) = p_R((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t \mathbf{P}_t)_{t=1}^T \mid \boldsymbol{\phi}), \tag{A.15}$$

for every $(\mathbf{P}_t)_{t=1}^T \in \mathcal{O}_n^T$. The result follows because Equation (A.15) is satisfied if and only if $p_R((\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi})$ does not depend on $(\mathbf{Q}_t)_{t=1}^T$. □

B Prior over $\boldsymbol{\phi}^{RC}$

This section summarizes the model constant parameters and describes our prior over $\boldsymbol{\phi}^{RC}$. For ease of exposition, we partition $\boldsymbol{\phi}^{RC}$ into fixed constant parameters $\boldsymbol{\phi}_F^{RC}$ and estimated constant parameters $\boldsymbol{\phi}_E^{RC}$, that is $\boldsymbol{\phi}^{RC} = (\boldsymbol{\phi}_F^{RC}, \boldsymbol{\phi}_E^{RC})$. The parameters in $\boldsymbol{\phi}_E^{RC}$ depend on hyperparameters, which we denote by $\boldsymbol{\psi}^{RC}$. Table A1 summarizes parameters and hyperparameters. Next, we turn to the details.

Table A1: Model Parameters

Fixed Constant Parameters: $\boldsymbol{\phi}_F^{RC}$	
$\mathbf{m}_{\boldsymbol{\beta}_1}$	Expected value of $\boldsymbol{\beta}_1$.
$\mathbf{V}_{\boldsymbol{\beta}_1}$	Variance of $\boldsymbol{\beta}_1$.
$\boldsymbol{m}_{\boldsymbol{\delta}_1}$	Expected value of $\boldsymbol{\delta}_1$.
$\boldsymbol{V}_{\boldsymbol{\delta}_1}$	Variance of $\boldsymbol{\delta}_1$.
$\boldsymbol{m}_{\boldsymbol{\gamma}_1}$	Expected value of $\boldsymbol{\gamma}_1$.
$\boldsymbol{V}_{\boldsymbol{\gamma}_1}$	Variance of $\boldsymbol{\gamma}_1$.
Estimated Constant Parameters: $\boldsymbol{\phi}_E^{RC}$	
$\mathbf{V}_{\boldsymbol{\beta}}$	Variance of the innovations to $\boldsymbol{\beta}_t$.
$\mathbf{V}_{\boldsymbol{\delta}}$	Variance of the innovations to $\boldsymbol{\delta}_t$.
$\mathbf{V}_{\boldsymbol{\gamma}}$	Variance of the innovations to $\boldsymbol{\gamma}_t$.
Hyperparameters: $\boldsymbol{\psi}^{RC}$	
$\bar{\mathbf{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}}$	Scale matrix of the prior for $\mathbf{V}_{\boldsymbol{\beta}}$.
$\bar{k}_{\mathbf{V}_{\boldsymbol{\beta}}}$	Scaling factor for the scale matrix of the prior for $\mathbf{V}_{\boldsymbol{\beta}}$.
$\bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}}$	Degrees of freedom of the prior for $\mathbf{V}_{\boldsymbol{\beta}}$.
$\bar{\nu}_{\delta,i}$	Shape parameter of the prior for $V_{\delta,i}$ for $i = 1, \ldots, n$.
$\Psi_{V_{\delta,i}}$	Scale parameter of the prior for $V_{\delta,i}$ for $i = 1, \ldots, n$.
$k_{w,i}$	Scaling factor for scale parameter of the prior for $V_{\delta,i}$ for $i = 1, \ldots, n$.
$\bar{\nu}_{\gamma,i}$	Shape parameter of the prior for $V_{\gamma,i}$ for $i = 1, \ldots, n_\gamma$.
$\Psi_{V_{\gamma,i}}$	Scale parameter of the prior for $V_{\gamma,i}$ for $i = 1, \ldots, n_\gamma$.
$k_{v,i}$	Scaling factor for scale parameter of the prior for $V_{\gamma,i}$ for $i = 1, \ldots, n_\gamma$.

The prior over $\boldsymbol{\phi}_F^{RC}$ will be Dirac. This assumption can be relaxed at the cost of increased computation time. To set a value for $\boldsymbol{\phi}_F^{RC}$, we set $\mathbf{m}_{\boldsymbol{\beta}_1}$ equal to the maximum likelihood estimate of a time-invariant VAR with the same variables and lags based on the first $T_0 = 40$

observations available in our sample. We denote such an estimate by $\hat{\mathbf{B}}$. We set $\mathbf{V}_{\boldsymbol{\beta}_1}$ equal to the unbiased estimator for the variance of $\hat{\mathbf{B}}$. To set the values for $\boldsymbol{m}_{\boldsymbol{\delta}_1}$, $\boldsymbol{m}_{\boldsymbol{\gamma}_1}$, $\boldsymbol{V}_{\boldsymbol{\delta}_1}$ and $\boldsymbol{V}_{\boldsymbol{\gamma}_1}$, first it will be useful to let $\hat{\boldsymbol{\Sigma}}$ denote the maximum likelihood estimate of the variance of the reduced-form residuals of the time-invariant VAR mentioned above, and second to define the following mapping between $\text{vech}(\hat{\boldsymbol{\Sigma}})$—where the vech operator stacks the elements on and below the main diagonal of a square matrix—and $(\boldsymbol{\delta}_1, \boldsymbol{\gamma}_1)$:

$$g\left(\text{vech}(\hat{\boldsymbol{\Sigma}})\right) = \left(\underbrace{2\log\left(\text{diag}(\hat{\mathbf{D}})\right)}_{\boldsymbol{\delta}_1}, \underbrace{\text{vecl}(\log \hat{\mathbf{C}})}_{\boldsymbol{\gamma}_1} \right),$$

where $\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1}\hat{\boldsymbol{\Sigma}}\hat{\mathbf{D}}^{-1}$, $\hat{\mathbf{D}} = \left(\text{diag}(\text{diag}(\hat{\boldsymbol{\Sigma}}))\right)^{\frac{1}{2}}$, $\hat{\boldsymbol{\Sigma}} = \left(\text{vec}(\mathbf{I}_n)' \otimes \mathbf{I}_n\right)\left(\mathbf{I}_n \otimes \left(\mathbf{D}_n \,\text{vech}(\hat{\boldsymbol{\Sigma}})\right)\right)$, and $\mathbf{D}_n$ is a $n^2 \times \frac{n(n+1)}{2}$ duplication matrix such that $\text{vec}(\hat{\boldsymbol{\Sigma}}) = \mathbf{D}_n \,\text{vech}(\hat{\boldsymbol{\Sigma}})$. By Proposition 3.4 of Lütkepohl (2007),

$$\sqrt{T}\left(\text{vech}(\hat{\boldsymbol{\Sigma}}) - \text{vech}(\boldsymbol{\Sigma})\right) \to \text{N}(\mathbf{0}, 2\,\mathbf{D}_n^{+}\left(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}\right)\mathbf{D}_n^{+\prime}) \tag{A.16}$$

where $\mathbf{D}_n^{+}$ is the Moore-Penrose generalized inverse of the duplication matrix $\mathbf{D}_n$. Then, by the Delta Method,

$$\sqrt{T}\left(g\left(\text{vech}(\hat{\boldsymbol{\Sigma}})\right) - g\left(\text{vech}(\boldsymbol{\Sigma})\right)\right) \to \text{N}(\mathbf{0}, \mathbf{D}_g(\boldsymbol{\Sigma})2\,\mathbf{D}_n^{+}\left(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}\right)\mathbf{D}_n^{+\prime}\mathbf{D}_g(\boldsymbol{\Sigma})')$$

where $\mathbf{D}_g\left(\boldsymbol{\Sigma}\right) = \frac{\partial g(\text{vech}(\boldsymbol{\Sigma}))}{\partial \text{vech}(\boldsymbol{\Sigma})}$. Let $\boldsymbol{V}_{g\left(\text{vech}(\hat{\boldsymbol{\Sigma}})\right)}(\boldsymbol{\Sigma}) = \frac{\mathbf{D}_g(\boldsymbol{\Sigma})2\,\mathbf{D}_n^{+}(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\,\mathbf{D}_n^{+\prime}\mathbf{D}_g(\boldsymbol{\Sigma})'}{T}$. Thus,

$$\boldsymbol{m}_{\boldsymbol{\delta}_1} = 2\log\left(\text{diag}(\hat{\mathbf{D}})\right) \quad , \quad \boldsymbol{m}_{\boldsymbol{\gamma}_1} = \text{vecl}(\log \hat{\mathbf{C}})$$

$$\boldsymbol{V}_{\boldsymbol{\delta}_1} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n,n_\gamma} \end{bmatrix} \boldsymbol{V}_{g\left(\text{vech}(\hat{\boldsymbol{\Sigma}})\right)}(\hat{\boldsymbol{\Sigma}}) \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0}_{n_\gamma,n} \end{bmatrix} \text{, and } \quad \boldsymbol{V}_{\boldsymbol{\gamma}_1} = \begin{bmatrix} \mathbf{0}_{n_\gamma,n} & \mathbf{I}_{n_\gamma} \end{bmatrix} \boldsymbol{V}_{g\left(\text{vech}(\hat{\boldsymbol{\Sigma}})\right)}(\hat{\boldsymbol{\Sigma}}) \begin{bmatrix} \mathbf{I}_{n_\gamma} \\ \mathbf{0}_{n,n_\gamma} \end{bmatrix}$$

where $\mathbf{I}_s$ is the identity matrix of dimension $s \times s$ and $\mathbf{0}_{s_1,s_2}$ is a matrix of zeros of dimension $s_1 \times s_2$.

The prior over $\boldsymbol{\phi}_E^{RC}$, i.e., $p(\boldsymbol{\phi}_E^{RC} \mid \boldsymbol{\phi}_F^{RC}, \boldsymbol{\psi}^{RC})$, is as follows:

$$
\begin{aligned}
\mathbf{V}_{\boldsymbol{\beta}} &\sim \text{IW}(\bar{k}^2_{\mathbf{V}_{\boldsymbol{\beta}}} \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}} \bar{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}}, \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}}) \\
V_{\delta,i} &\sim \text{IG}\left(\frac{\bar{\nu}_{\delta,i}}{2}, \frac{k^2_{w,i}\bar{\nu}_{\delta,i}\Psi_{V_{\delta,i}}}{2}\right), \text{for } i = 1, \ldots, n \\
V_{\gamma,i} &\sim \text{IG}\left(\frac{\bar{\nu}_{\gamma,i}}{2}, \frac{k^2_{v,i}\bar{\nu}_{\gamma,i}\Psi_{V_{\gamma,i}}}{2}\right), \text{for } i = 1, \ldots, n_\gamma
\end{aligned}
$$

The values for the hyperparameters ($\boldsymbol{\psi}_E^{RC}$) are: $\bar{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}} = \boldsymbol{V}_{\boldsymbol{\beta}_1}$, $\bar{k}_{\mathbf{V}_{\boldsymbol{\beta}}} = 0.01$, $\bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}} = nm + 2$, $\bar{\nu}_{\delta,i} = 3$, $\Psi_{V_{\delta,i}} = 4 \times e'_{i,n} \boldsymbol{V}_{\boldsymbol{\delta}_1} e_{i,n}$, $k_{w,i} = 0.1$, for $i = 1 \ldots, n$, and $\bar{\nu}_{\gamma,i} = 6$, $\Psi_{V_{\gamma,i}} = 4 \times e'_{i,n_\gamma} \boldsymbol{V}_{\boldsymbol{\gamma}_1} e_{i,n_\gamma}$, $k_{v,i} = 0.1$, for $i = 1 \ldots, n_\gamma$, where $e_{i,x}$ denotes the i-th column of an identity matrix of dimension x.

C Efficient Algorithm

We use the particle Gibbs with ancestor sampling (PGAS) developed by Lindsten, Jordan, and Schon (2014). The PGAS iterates over multiple blocks to draw from $p((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$. It differs from a traditional Gibbs sampling algorithm as we draw $(\mathbf{A}_t^i)_{t=1}^T$ from its conditional posterior distribution using the conditional sequential Monte Carlo method. This is because we are conditioning the sign restrictions; therefore, its conditional posterior distribution is no longer Gaussian. We use an off-the-shelf implementation of the PGAS for state-space models presented in Algorithm 3 of Lindsten, Jordan and Schon (2014).

Steps for drawing $(\mathbf{F}_t^i)_{t=1}^T$ and $\boldsymbol{\phi}^{i,RC}$ are based on their exact conditional posterior distribution, and therefore they are fairly standard. To simplify the exposition, we assume that $\boldsymbol{S}_t(\mathbf{A}_t, \mathbf{F}_t) = \boldsymbol{S}_t(\mathbf{A}_t)$. For the general case where $\mathbf{F}_t$ enters the sign restrictions, one can draw $\mathbf{F}_t$ from its conditional posterior distribution using the conditional sequential Monte Carlo method as in the step for drawing $\mathbf{A}_t$. For ease of exposition, we first reproduce Algorithm 3 and then discuss each step's details.

Algorithm C.1. *The following algorithm draws from* $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ *conditional on the sign restrictions.*

1. *Let* $N_1, N_2 > 1$*, set* $i = 1$*, and initialize* $((\mathbf{A}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC})$.
2. *Draw* $(\mathbf{F}_t^i)_{t=1}^T$ *from a distribution that is proportional to* $p_S^{RC}((\mathbf{F}_t)_{t=1}^T \mid (\mathbf{A}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}, (\mathbf{y}_t)_{t=1}^T)$.
3. *Draw* $(\mathbf{A}_t^i)_{t=1}^T$ *from* $p_S^{RC}((\mathbf{A}_t)_{t=1}^T \mid (\mathbf{F}_t^i)_{t=1}^T, (\mathbf{A}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}, (\mathbf{y}_t)_{t=1}^T)$ *conditional on the sign restrictions.*
4. *Draw* $(\boldsymbol{\phi}^{i,RC})$ *from* $p_S^{RC}(\boldsymbol{\phi}^{RC} \mid (\mathbf{A}_t^i, \mathbf{F}_t^i)_{t=1}^T, (\mathbf{y}_t)_{t=1}^T)$.
5. *If* $i < N_1 + N_2$*, let* $i = i + 1$ *and return to Step 2.*

C.1 Sampling from $p\left((\boldsymbol{F}_t)_{t=1}^T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$

In order to draw from $p\left((\boldsymbol{F}_t)_{t=1}^T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$, we will draw from

$$p\left((\boldsymbol{\beta}_t)_{t=1}^T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$$

with $\boldsymbol{\beta}_t = \text{vec}(\boldsymbol{B}_t)$ and transform the draws to $(\boldsymbol{F}_t)_{t=1}^T$ exploiting the linear mapping between $\boldsymbol{F}_t$ and $\boldsymbol{B}_t$, given $\boldsymbol{A}_t$.

Let $\boldsymbol{\beta}_{T|T} = \mathbb{E}\left(\boldsymbol{\beta}_T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$, $\mathbf{V}_{T|T} = \text{Var}\left(\boldsymbol{\beta}_T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$, $\boldsymbol{\beta}_{t|t+1} = \mathbb{E}\left(\boldsymbol{\beta}_t \mid \boldsymbol{\beta}_{t+1}, (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$, and $\mathbf{V}_{t|t+1} = \text{Var}\left(\boldsymbol{\beta}_t \mid \boldsymbol{\beta}_{t+1}, (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$ for $t = T-1, \ldots, 1$. Following Carter and Kohn (1994),

$$p\left((\boldsymbol{\beta}_t)_{t=1}^T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right) = p\left((\boldsymbol{\beta}_t)_{t=1}^T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right)$$

and:

$$p\left((\boldsymbol{\beta}_t)_{t=1}^T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_t)_{t=1}^T\right) = p\left((\boldsymbol{\beta}_t)_{t=1}^T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right) \prod_{t=1}^{T-1} p\left(\boldsymbol{\beta}_t \mid \boldsymbol{\beta}_{t+1}, (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right) \tag{A.17}$$

where, for $t = T-1, \ldots, 1$,

$$\begin{aligned} p\left(\boldsymbol{\beta}_T \mid (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_t)_{t=1}^T\right) &= \mathrm{N}_{(\boldsymbol{\beta}_{T|T}, \boldsymbol{V}_{T|T})}(\boldsymbol{\beta}_T) \\ p\left(\boldsymbol{\beta}_t \mid \boldsymbol{\beta}_{t+1}, (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_s)_{s=1}^t\right) &= \mathrm{N}_{(\boldsymbol{\beta}_{t|t+1}, \boldsymbol{V}_{t|t+1})}(\boldsymbol{\beta}_t). \end{aligned}$$

Conditional on $(\boldsymbol{A}_t)_{t=1}^T$, the draws of $\boldsymbol{\beta}_t$ can be mapped to draws of $\mathbf{F}_t$ using $\mathbf{F}_t = \mathbf{B}_t \mathbf{A}_t$, for $t = 1, \ldots, T$.

C.2 Sampling from $p_S^{RC}((\mathbf{A}_t)_{t=1}^T \mid (\mathbf{F}_t)_{t=1}^T, (\mathbf{A}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{RC}, (\mathbf{y}_t)_{t=1}^T)$

Let $\mathbf{1}\{\boldsymbol{S}_t(\mathbf{A}_t) > \mathbf{0}\}$ be an indicator function that equals 1 if the sign restrictions are satisfied at time t, and 0 otherwise. In addition, let superscript i index the Gibbs sampler draws and let subscript j index the particles and the corresponding weights used to approximate $p_S^{RC}((\mathbf{A}_t)_{t=1}^T \mid (\mathbf{F}_t^i)_{t=1}^T, (\mathbf{A}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}, (\mathbf{y}_t)_{t=1}^T)$.

Algorithm C.2. *The following algorithm draws from $p_S^{RC}((\mathbf{A}_t)_{t=1}^T \mid (\mathbf{F}_t^i)_{t=1}^T, (\mathbf{A}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i,RC}, (\mathbf{y}_t)_{t=1}^T)$ conditional on the sign restrictions. Set $J > 1$ and $t = 1$ and enter the following steps:*

1. *Draw $\mathbf{A}_{t,j} \sim p_S^{RC}(\mathbf{A}_1 \mid (\mathbf{F}_t^i)_{t=1}^T, \boldsymbol{\phi}^{i,RC})$ for $j = 1, ..., J-1$.*
2. *Set $\mathbf{A}_{t,J} = \mathbf{A}_t^{(i-1)}$.*
3. *Set $w_{t,j} \propto p(\mathbf{y}_t \mid \mathbf{A}_{t,j}, \mathbf{F}_t^i, \mathbf{x}_t)\mathbf{1}\{\boldsymbol{S}_t(\mathbf{A}_{t,j}) > 0\}$ for $j = 1, ..., J$.*
4. *Set $t = t + 1$.*
5. *Generate $\{(\widetilde{\mathbf{A}}_{s,j})_{s=1}^{t-1}\}_{j=1}^{J-1}$ by sampling $J-1$ times with replacement from $\{(\mathbf{A}_{s,j})_{s=1}^{t-1}\}_{j=1}^{J}$ with probabilities proportional to the importance weights $\{w_{t-1,j}\}_{j=1}^{J}$.*
6. *Draw J' with $P(J' = j) \propto w_{t-1,j} p(\mathbf{A}_t^{i-1} \mid \mathbf{A}_{t-1,j})$. And, set $(\widetilde{\mathbf{A}_{s,J}})_{s=1}^{t-1} = (\mathbf{A}_{s,J'})_{s=1}^{t-1}$.*
7. *Simulate $\mathbf{A}_{t,j} \sim p_S^{RC}(\mathbf{A}_t \mid \widetilde{\mathbf{A}}_{t-1}, \boldsymbol{\phi}^{RC})$ for $j = 1, ..., J-1$.*
8. *Set $\mathbf{A}_{t,J} = \mathbf{A}_t^{i-1}$.*
9. *Set $(\mathbf{A}_{s,j})_{s=1}^{t} = ((\widetilde{\mathbf{A}}_{s,j})_{s=1}^{t-1}, \mathbf{A}_{t,j})$ for $j = 1, ..., J$.*

10. Set $w_{t,j} = p(\mathbf{y}_t \mid \mathbf{A}_{t,j}, \mathbf{F}_t^i, \mathbf{x}_t)\mathbf{1}\{\boldsymbol{S}_t(\mathbf{A}_{t,j}) > 0\}$ *for* $j = 1, ..., J$.

11. If $t < T$, *return to Step 4. Otherwise, go to the next step.*

12. Draw k *with* $P(k = j) \propto w_{T,j}$.

13. Set $(\mathbf{A}_t^i)_{t=1}^T = \mathbf{A}_{1:T,k}$ *and exit the algorithm.*

The above algorithm can be understood sequentially constructing the important sampling weight, $w_{t,j}$, for the path of the j-th particle, $(\mathbf{A}_{s,j})_{s=1}^t = \{\mathbf{A}_{1,j}, \mathbf{A}_{2,j}, ..., \mathbf{A}_{t,j}\}$. Then, the final particle system $\{w_{T,j}, (\mathbf{A}_{s,j})_{s=1}^T\}_{j=1}^J$ approximates the desired target distribution and Steps 12 and 13 generate a draw from $p_S^{RC}((\mathbf{A}_t)_{t=1}^T \mid (\mathbf{F}_t^i)_{t=1}^T, (\mathbf{A}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}, (\mathbf{y}_t)_{t=1}^T)$.

Next, we note several important points about the algorithm. First, this algorithm is almost identical to the typical particle filter algorithm except for Steps 2, 8, and 6. Steps 2 and 8 make sure that the algorithm conditions on the previous draw, $(\mathbf{A}_t^{i-1})_{t=1}^T$, by forcing the last particle to be the same as the previous draw. Resampling for the last particle needs to be modified accordingly as we need to resample $\{(\mathbf{A}_{s,j})_{s=1}^{t-1}\}_{j=1}^J$ conditional on $\mathbf{A}_t^{i-1}$, which is done in Step 6. Lindsten, Jordan and Schon (2014) prove that these extra steps ensure that Algorithm 3 draws from the desired joint posterior distribution.

Second, let $(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t) = g_t^{RC}(\boldsymbol{\Sigma}_t)$, $\boldsymbol{\Sigma}_t = (\mathbf{A}_t \mathbf{A}_t')^{-1}$, and $f_h^a(\mathbf{A}_t) = (\boldsymbol{\Sigma}_t, \boldsymbol{Q}_t)$ with $\boldsymbol{Q}_t = h(\boldsymbol{\Sigma}_t)\mathbf{A}_t$. Let $t = 1$ and notice that if we combine the prior over $(\boldsymbol{\delta}_1, \boldsymbol{\gamma}_1)$ with the uniform prior over $\boldsymbol{Q}_1$, the implied prior over $\mathbf{A}_1$ is

$$\begin{aligned} p_S^{RC}(\mathbf{A}_1 \mid (\mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}) &= p_S^{RC}(\mathbf{A}_1 \mid \boldsymbol{\phi}^{i-1,RC}) \\ &= 2^{\frac{n(n+1)}{2}} |\det(\mathbf{A}_1)|^{-(2n+m+1)} v_{g_1^{RC}}(\boldsymbol{\Sigma}_1) \frac{p^{RC}((\boldsymbol{\delta}_1, \boldsymbol{\gamma}_1) \mid \boldsymbol{\phi}^{i-1,RC})}{\int_{\mathbb{O}_n} 1 d_{\mathcal{O}_n} \mathbf{Q}_1}. \end{aligned}$$

Hence, in Step 1 we can draw from $p_S^{RC}(\mathbf{A}_1 \mid \boldsymbol{\phi}^{RC})$ drawing $(\boldsymbol{\delta}_1, \boldsymbol{\gamma}_1)$ from $p^{RC}((\boldsymbol{\delta}_1, \boldsymbol{\gamma}_1) \mid \boldsymbol{\phi}^{i-1,RC})$ and $\boldsymbol{Q}_1$ from the uniform distribution over $\mathbb{O}_n$, and then transforming the draws, first into $\boldsymbol{\Sigma}_1$ and then into $\mathbf{A}_1$.

Similarly,

$$p_S^{RC}(\mathbf{A}_t \mid \mathbf{A}_{t-1}, \boldsymbol{\phi}^{i-1,RC}) = 2^{\frac{n(n+1)}{2}} |\det(\mathbf{A}_t)|^{-(2n+m+1)} v_{g_t^{RC}}(\boldsymbol{\Sigma}_t) \frac{p^{RC}((\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t) \mid \boldsymbol{\phi}^{i-1,RC})}{\int_{\mathbb{O}_n} 1 d_{\mathcal{O}_n} \mathbf{Q}_t}, \quad \text{(A.18)}$$

for $t = 2, \ldots, T$. Hence, in Step 7 we can draw from $p_S^{RC}(\mathbf{A}_t \mid \mathbf{A}_{t-1}, \boldsymbol{\phi}^{i-1,RC})$ by drawing $(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)$ from $p^{RC}((\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t) \mid g^R((\mathbf{A}_{t-1}\,\mathbf{A}_{t-1}')^{-1}), \boldsymbol{\phi}^{i-1,RC})$ and $\boldsymbol{Q}_t$ from the uniform distribution over $\mathbb{O}_n$, and the transforming the draws into $\mathbf{A}_t$. In some cases, it is more efficient to combine the mappings g_t^{RC} and f_h^a to form $f_h^{a\to\boldsymbol{\delta}\boldsymbol{\gamma}\boldsymbol{Q}}(\mathbf{A}_t) = (\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \boldsymbol{Q}_t)$. This mapping can be used to sample from $p^{RC}(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t \mid \boldsymbol{\delta}_{t-1}, \boldsymbol{\gamma}_{t-1}, \mathbf{Q}_{t-1}, \boldsymbol{\phi}^{i-1,RC})$, which is a density defined as follows:

$$v_{(f_h^{a\to\boldsymbol{\delta}\boldsymbol{\gamma}\boldsymbol{Q}})^{-1}(\boldsymbol{\delta}_t,\boldsymbol{\gamma}_t,\boldsymbol{Q}_t)} p_S^{RC}((f_h^{a\to\boldsymbol{\delta}\boldsymbol{\gamma}\boldsymbol{Q}})^{-1}(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \boldsymbol{Q}_t) \mid (f_h^{a\to\boldsymbol{\delta}\boldsymbol{\gamma}\boldsymbol{Q}})^{-1}(\boldsymbol{\delta}_{t-1}, \boldsymbol{\gamma}_{t-1}, \boldsymbol{Q}_{t-1}), \boldsymbol{\phi}^{i-1,RC})$$

by means of a Gibbs Sampler consisting of sampling from $p^{RC}(\mathbf{Q}_t \mid \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \boldsymbol{\delta}_{t-1}, \boldsymbol{\gamma}_{t-1}, \mathbf{Q}_{t-1}, \boldsymbol{\phi}^{i-1,RC})$ and then from $p^{RC}(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t \mid \mathbf{Q}_t, \boldsymbol{\delta}_{t-1}, \boldsymbol{\gamma}_{t-1}, \mathbf{Q}_{t-1}, \boldsymbol{\phi}^{i-1,RC})$.

Finally, the indicator $\mathbf{1}\{\boldsymbol{S}_t(\mathbf{A}_{t,j}) > 0\}$ in Step 3 and 10 discards the draws that do not satisfy the sign restrictions. It is important to notice that this is done period by period, unlike the simple algorithm such as Algorithm 1.

C.3 Sampling from $p_S^{RC}(\boldsymbol{\phi}^{RC} \mid (\mathbf{A}_t^i, \mathbf{F}_t^i)_{t=1}^T, (\mathbf{y}_t)_{t=1}^T)$

Recall that $\boldsymbol{\phi}^{RC}$ consists of several hyperparameters in the model, $\mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}$, and $\mathbf{V}_{\boldsymbol{\gamma}}$.

C.3.1 Sampling from $p(\mathbf{V}_{\boldsymbol{\beta}} \mid (\boldsymbol{F}_t)_{t=1}^T, (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\delta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_t)_{t=1}^T)$

Our prior distribution over $\mathbf{V}_{\boldsymbol{\beta}}$ is $\text{IW}(\bar{k}_{\mathbf{V}_{\boldsymbol{\beta}}}^2 \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}} \bar{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}}, \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}})$, then the posterior is:

$$\begin{aligned} p(\mathbf{V}_{\boldsymbol{\beta}} \mid (\boldsymbol{F}_t)_{t=1}^T, (\boldsymbol{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\alpha}}, (\mathbf{y}_t)_{t=1}^T) &= p(\mathbf{V}_{\boldsymbol{\beta}} | (\boldsymbol{\beta}_t)_{t=1}^T, (\boldsymbol{A}_t)_{t=1}^T) \\ &\propto p((\boldsymbol{\beta}_t)_{t=1}^T \mid \mathbf{V}_{\boldsymbol{\beta}}) p(\mathbf{V}_{\boldsymbol{\beta}}) \propto \text{IW}_{(\tilde{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}}, \tilde{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}})}(\mathbf{V}_{\boldsymbol{\beta}}), \end{aligned}$$

where $\tilde{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}} = (T-1) + \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}}$ and $\tilde{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}} = (\mathcal{B}_{2:T} - \mathcal{B}_{1:T-1})'(\mathcal{B}_{2:T} - \mathcal{B}_{1:T-1}) + \bar{k}_{\mathbf{V}_{\boldsymbol{\beta}}}^2 \bar{\nu}_{\mathbf{V}_{\boldsymbol{\beta}}} \bar{\boldsymbol{\Psi}}_{\mathbf{V}_{\boldsymbol{\beta}}}$, where $\mathcal{B}_{s:r} = (\boldsymbol{\beta}_s, \ldots, \boldsymbol{\beta}_r)'$ for $1 < s < r \leq T$.

C.3.2 Sampling from $p(\mathbf{V}_{\boldsymbol{\delta}} \mid (\mathbf{F}_t)_{t=1}^T, (\mathbf{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_t)_{t=1}^T)$

Our prior distribution over V_{δ_i} is $\text{IG}\left(\frac{\bar{\nu}_{\delta,i}}{2}, \frac{\bar{\psi}_{\delta,i}}{2}\right)$ for $i = 1, \ldots, n$. Then, the posterior is:

$$p(V_{\delta_i} \mid (\mathbf{F}_t)_{t=1}^T, (\mathbf{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_t)_{t=1}^T) = p(V_{\delta_i} \mid \mathcal{D}_T) \propto p(\mathcal{D}_T \mid V_{\delta_i}) p(V_{\delta_i}) \propto \text{IG}_{\left(\frac{\tilde{\nu}_{\delta,i}}{2}, \frac{\tilde{\psi}_{\delta,i}}{2}\right)}(V_{\delta_i}),$$

where $\tilde{\nu}_{\delta_i} = (T-1) + \bar{v}_{\delta,i}$ and $\tilde{\psi}_{\delta.i} = \bar{\psi}_{\delta,i} + \sum_{t=2}^T \delta_{i,t}^2$, where $\delta_{i,t}$ denotes the i-th entry of $\boldsymbol{\delta}_t$.

C.3.3 Sampling from $p(\mathbf{V}_{\boldsymbol{\gamma}} \mid (\mathbf{F}_t)_{t=1}^T, (\mathbf{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\delta}}, (\mathbf{y}_t)_{t=1}^T)$

Our prior distribution over V_{γ_i} is $\text{IG}\left(\frac{\bar{\nu}_{\gamma,i}}{2}, \frac{\bar{\psi}_{\gamma,i}}{2}\right)$ for $i = 1, \ldots, n_\gamma$. Then, the posterior is:

$$p(V_{\gamma_i} \mid (\mathbf{F}_t)_{t=1}^T, (\mathbf{A}_t)_{t=1}^T, \mathbf{V}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\gamma}}, (\mathbf{y}_t)_{t=1}^T) = p(V_{\gamma_i} \mid (\mathbf{y}_t)_{t=1}^T) \propto p((\mathbf{y}_t)_{t=1}^T \mid V_{\gamma_i}) p(V_{\gamma_i}) \propto \text{IG}_{\left(\frac{\tilde{\nu}_{\gamma,i}}{2}, \frac{\tilde{\psi}_{\gamma,i}}{2}\right)}(V_{\gamma_i}),$$

where $\tilde{\nu}_{\gamma_i} = (T-1) + \bar{\nu}_{\delta,i}$ and $\tilde{\psi}_{\gamma.i} = \bar{\psi}_{\gamma,i} + \sum_{t=2}^T \gamma_{i,t}^2$, where $\gamma_{i,t}$ denotes the i-th entry of $\boldsymbol{\gamma}_t$.

D Data References

The U.S. Bureau of Economic Analysis, Real Gross Domestic Product [GDPC1], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GDPC1, September 9, 2023. U.S. Bureau of Economic Analysis, Personal Consumption Expenditures Excluding Food and Energy (Chain-Type Price Index) [PCEPILFE], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PCEPILFE, September 9, 2023. Board of Governors of the Federal Reserve System (US), Federal Funds Effective Rate [FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FEDFUNDS, September 9, 2023. Board of Governors of the Federal Reserve System (US), M2 [M2SL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/M2SL, September 9, 2023. Federal Reserve Bank of St. Louis, Moody's Seasoned Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity [BAA10YM], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/BAA10YM, September 9, 2023.