Inference Based on Time-Varying SVARs Identified with Sign Restrictions

Jonas E. Arias
Federal Reserve Bank of Philadelphia

Juan F. Rubio-Ramírez*
Emory University
Federal Reserve Bank of Atlanta

Minchul Shin
Federal Reserve Bank of Philadelphia

Daniel F. Waggoner
Emory University
Federal Reserve Bank of Atlanta, Emeritus

January 11, 2026

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
Keywords: time-varying parameters, structural vector autoregressions, identification.

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.

*Corresponding author: Juan F. Rubio-Ramírez <juan.rubio-ramirez@emory.edu>, Economics Department, Emory University, Rich Memorial Building, Room 306, Atlanta, Georgia 30322-2240. We thank Mark Bognanni, Frank Schorfheide, Christian Wolf, and Jonathan Wright for their helpful comments.

1 Introduction

Structural vector autoregressions (SVARs) featuring time-varying parameters are widely used in empirical macroeconomics to explore a variety of questions, such as the economic impacts of policy shocks, oil price shocks, the interaction between financial markets and economic activity, and the role of monetary policy during the Great Inflation.[1] This paper makes two main contributions to the Bayesian approach for inference based on time-varying SVARs using the class of rotation-invariant models introduced by Bognanni (2018) as a foundation. Members of this class are defined by a measurement equation linking observed variables to a sequence of time-varying structural parameters and a rotation-invariant prior that remains unchanged under orthogonal transformations of these sequences. A rotation-invariant prior is both a necessary and sufficient condition for ensuring that the posterior over the sequences of time-varying structural parameters is identical across observationally equivalent sequences, thereby necessitating identifying restrictions for causal inference. To address this identification problem, we employ traditional sign restrictions, typically derived from economic theory and institutional knowledge. Our methods also allow us to consider time-varying sign restrictions, i.e., identification restrictions that vary across the sample. Importantly, working with rotation-invariant priors ensures that only the sign restrictions influence the posterior distribution over observationally equivalent sequences of time-varying structural parameters. Consequently, within this framework, researchers can separate reduced-form estimation from causal inference, thereby preserving the benefits that have made sign restrictions popular in SVARs with constant parameters.

The first contribution of this paper is to characterize the class of rotation-invariant time-varying SVARs models in terms of priors over sequences of time-varying orthogonal reduced-form parameters. Specifically, expanding the work of Uhlig (2005) and Rubio-Ramírez et al. (2010), we define an invertible mapping between the sequences of time-varying structural parameters and sequences of time-varying orthogonal reduced-form parameters. With this mapping, we demonstrate that any prior over the sequences of time-varying reduced-form

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

parameters, combined with an independent uniform prior over the sequences of orthogonal matrices, defines an element of the class of rotation-invariant time-varying SVARs models. This result is a crucial if-and-only-if condition that delineates the scope of our methodology: any of the commonly used priors for the time-varying reduced-form parameters—e.g., Primiceri (2005); Cogley and Sargent (2005) and the dynamic linear models with discounted Wishart stochastic volatility inspired by Uhlig (1994, 1997) and documented in Bognanni (2018)—can be adapted for structural analysis when combined with an independent uniform prior over the sequences of orthogonal matrices. This underscores that the class of rotation-invariant time-varying SVARs models is extensive, enhancing its practical utility. However, as emphasized by Baumeister and Hamilton (2015); Giacomini and Kitagawa (2021), it is crucial to recognize that sign restrictions yield identified sets, and the outcomes are influenced not only by these restrictions but also by the uniform prior imposed on the sequence of orthogonal rotation matrices.

The second contribution is to introduce a Gibbs Sampler algorithm that can be used to draw from the posterior distribution of any element of the class of rotation-invariant time-varying SVARs models conditional on the time-varying sign restrictions. The algorithm exploits the slice elliptical sampling developed by Murray et al. (2010) and is the first to draw from the class of rotation-invariant time-varying SVARs. Previous attempts to draw from the posterior of particular elements of the class (e.g., Baumeister and Peersman, 2013; Bognanni, 2018; Debortoli et al., 2020) consider a uniform distribution over the sequences of orthogonal matrices conditional on the reduced-form parameters and the sign restrictions. Although this approach simplifies posterior sampling, once a prior over the sequences of time-varying reduced-form parameters is specified, this prior and the prior over the sequences of orthogonal matrices are not independent. Consequently, the draws are from a model that does not belong to the class of rotation-invariant time-varying SVARs, and the posterior distribution over observationally equivalent sequences of time-varying structural parameters is not solely determined by the sign restrictions. When describing the algorithm, we use a specific prior over the time-varying reduced-form parameters that relies on Archakov and Hansen's (2021) novel parameterization of correlation matrices. Hence, we refer to it as the Random Correlations prior, and we demonstrate that it defines an element of the class

that we will call the Random Correlations SVAR (RC-SVAR). Our rationale for using the RC-SVAR is motivated by the insights in Giannone et al. (2015), which highlight that a natural way to assess the impact of priors is by evaluating their implied out-of-sample forecasting performance. Arias et al. (2023) show that for the empirical applications typically considered, the Random Correlations prior generally implies a higher log-predictive score than most alternative models. While the algorithm will be described in terms of RC-SVARs, it can 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 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 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? Allowing for time variation in both the structural parameters and the sign restrictions 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 et al., 2000). In addition, certain structural relations in the economy may have changed. Our econometric approach provides a helpful setting to discipline inference about 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.

To answer the first question, 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 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 simulations 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 0.5 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 1.7 percent. Turning to the third question, our model estimates support the view that the Federal Reserve was behind the curve (see, e.g., Summers, 2021) in 2021. Nevertheless, 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.

Additionally, the Online Appendix presents a series of robustness checks to further validate our findings. Specifically, we examine the effects of distinguishing between two types of monetary policy shocks: one that primarily influences the short end of the yield curve and another that impacts the long end. This distinction allows us to explore the nuanced effects of monetary policy across different maturities. Furthermore, we assess the sensitivity of our results to the inclusion of both time-varying sign restrictions and time-varying parameters, emphasizing that the consideration of both elements is critical in the comprehensive analysis

of time-varying SVARs models. Our analysis underscores the importance of accounting for these factors to ensure that our conclusions are robust and reflective of the underlying economic dynamics.

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. The first element is 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, $(\mathbf{y}_0, \cdots, \mathbf{y}_{1-p})$, can have any distribution with full support. 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 $p_S\left(\mathbf{A}_t, \mathbf{F}_t \mid (\mathbf{A}_j, \mathbf{F}_j)_{j=1}^{t-1}, \boldsymbol{\phi}\right) = p_S\left(\mathbf{A}_t, \mathbf{F}_t \mid \mathbf{A}_{t-1}, \mathbf{F}_{t-1}, \boldsymbol{\phi}\right)$. Importantly, we assume that the law of motion satisfies:

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

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

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]

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] Equation (3) must hold for every $\boldsymbol{\phi}$ in the support of the marginal prior. We say that $((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi})$ 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 parameter 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* $(\tilde{\mathbf{A}}_t, \tilde{\mathbf{F}}_t)_{t=1}^T = (\mathbf{A}_t \mathbf{Q}_t, \mathbf{F}_t \mathbf{Q}_t)_{t=1}^T$.

See Appendix A for the proof. A similar (but not identical) proposition can be found

[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.

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

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

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 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) and rotation-invariant priors to those consistent with Equation (3).[8] 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 will be discussed in Section 5.

2.1 Heteroskedastic Structural Shocks

In this section, we now discuss models in which heteroskedastic structural shocks are the only source of time variation in the model. In particular, we now show how the condition 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:

$$\mathbf{y}_t' \mathbf{A} = \mathbf{x}_t' \mathbf{F} + \tilde{\boldsymbol{\varepsilon}}_t' \text{ with } \tilde{\boldsymbol{\varepsilon}}_t \sim \mathrm{N}\left(\mathbf{0}_n, \boldsymbol{\Psi}_t\right) \text{ for } 1 \leq t \leq T, \tag{5}$$

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. Lütkepohl and Netšunajev (2017) describe several ways to model the law of motion of $\boldsymbol{\Psi}_t$.[9] With these ingredients at hand, we can now formally show the following result.

[8] Multiplication by an $n \times n$ orthogonal matrix is either a rotation of $\mathbb{R}^n$ or a rotation of $\mathbb{R}^n$ times a reflection of $\mathbb{R}^n$, however, it is embedded in the literature to refer to orthogonal matrices as rotations. So, we take rotation invariant to mean invariant to multiplication by any orthogonal matrix.

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

Proposition 2. *Models with heteroskedastic structural shocks as defined in Equation (5) do not belong to the class of rotation-invariant time-varying SVARs models.*

See Appendix A for the proof. Similar arguments show that the restriction given by Equation (3) excludes the models described in Sentana and Fiorentini (2001); Rigobon (2003); Lanne and Lütkepohl (2008); Lanne et al. (2010); Brunnermeier et al. (2021). Furthermore, a model with time-varying $\mathbf{F}_t$ and $\mathbf{\Psi}_t$ but time-invariant $\mathbf{A}$ is also excluded from the class by the same reasoning, as we can still write $\mathbf{A}_t = \mathbf{A}\,\mathbf{\Psi}_t^{-1}$ in the proof above, and the argument follows.[10] 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. It is important to mention that Bognanni (2018) presents a concrete instance of a fully specified probability model that belongs to the class; we will give details about this model in Section 4.

While heteroskedasticity of the structural shocks can be exploited for identification, these shocks may not have a meaningful structural economic interpretation Herwartz and Lütkepohl (2014). 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.[11] This parameterization is characterized by a sequence of time-varying reduced-form parameters $(\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T$ and a sequence of time-varying orthogonal matrices $(\mathbf{Q}_t)_{t=1}^T$, and the measurement

[10] This is relevant for those willing to give a structural interpretation to Cogley and Sargent (2005).
[11] See Arias et al. (2018) for a definition in the constant parameter case.

equation can be written as follows:

$$\mathbf{y}_t' = \mathbf{x}_t' \mathbf{B}_t + \boldsymbol{\varepsilon}_t' \mathbf{Q}_t{}' h(\boldsymbol{\Sigma}_\mathbf{t}) \text{ for } 1 \leq t \leq T, \tag{6}$$

where the $n \times n$ matrix $h(\boldsymbol{\Sigma})$ is any decomposition of the variance-covariance matrix $\boldsymbol{\Sigma}$ satisfying $h(\boldsymbol{\Sigma})' h(\boldsymbol{\Sigma}) = \boldsymbol{\Sigma}$. We will take h to be the upper triangular 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:

$$f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) = \Big(\underbrace{\mathbf{F}_t \mathbf{A}_t^{-1}}_{\mathbf{B}_t}, \underbrace{(\mathbf{A}_t \mathbf{A}_t')^{-1}}_{\boldsymbol{\Sigma}_t}, \underbrace{h((\mathbf{A}_t \mathbf{A}_t')^{-1}) \mathbf{A}_t}_{\mathbf{Q}_t} \Big)_{t=1}^T. \tag{7}$$

This function is invertible and its inverse is given by:

$$f_h^{-1}((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T) = \Big(\underbrace{h(\boldsymbol{\Sigma}_t)^{-1} \mathbf{Q}_t}_{\mathbf{A}_t}, \underbrace{\mathbf{B}_t h(\boldsymbol{\Sigma}_t)^{-1} \mathbf{Q}_t}_{\mathbf{F}_t} \Big)_{t=1}^T. \tag{8}$$

Let $p_{OR}((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi})$ denote a prior over the time-varying orthogonal reduced-form parameters, conditional on $\boldsymbol{\phi}$.[12] The functions defined by Equations (7) and (8) 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 et al. (2018), the volume element of the mapping given by Equation (7) is $v_{f_h}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) = \prod_{t=1}^T 2^{\frac{n(n+1)}{2}} \mid \det(\mathbf{A}_t) \mid^{-(2n+m+1)}$ and the volume element of the mapping given by Equation (8) is $v_{f_h^{-1}}((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T) = \prod_{t=1}^T 2^{-\frac{n(n+1)}{2}} \mid \det(\boldsymbol{\Sigma}_t) \mid^{-\frac{2n+m+1}{2}}$. Thus, the prior over the time-varying orthogonal reduced-form parameters induced by p_S is:

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

[12] 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."

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

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

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

Proposition 3. *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* $(\mathbf{Q}_t)_{t=1}^T$.

See Appendix A for the proof. 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 $(\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T$ 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 $\mathcal{O}_n^T$.[13] 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 satisfy the restriction given by Equation (3).[14] In other words, any prior over the time-varying orthogonal reduced-form parameters that can be written as:

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

where $v(\mathcal{O}_n)$ is the volume of $\mathcal{O}_n$ with respect to the volume measure over $\mathcal{O}_n$ and $p_R((\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi})$ denotes the prior over the time-varying reduced-form parameters. Hence, every prior over $(\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T$ corresponds to an element of the class of rotation-invariant time-varying SVARs

[13] The proposition generalizes Proposition 4 in Bognanni (2018). The Haar measure over $\mathcal{O}_n$, which is only defined up to a scale factor, is any measure that is invariant with respect to rigid transformations, rotations, and reflections in this case. The volume measure is a Haar measure but with the scale determined by Lebesgue measure on $\mathbf{R}^{n^2}$, the set of all $n \times n$ matrices. The volume measure is the natural measure over orthogonal matrices in the sense that the Lebesgue measure is the natural measure over Euclidean spaces.

[14] Furthermore, 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.

models, and the prior over the time-varying structural parameters induced by p_R is:

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

where $\pi(\cdot)$ denotes the projection of $(\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T$ onto $(\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T$. 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 et al. (2023). This prior defines an element of the class of rotation-invariant time-varying SVARs models. These results depend only on the fact that the volume element does not depend on the sequence of orthogonal matrices. Hence, Proposition 3 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 $\mathbf{D}_t = \operatorname{diag}(\operatorname{diag}(\boldsymbol{\Sigma}_t)^{\frac{1}{2}})$ 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.[15] We can map $\boldsymbol{\Sigma}_t$ to $(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t) \in \mathbb{R}^n \times \mathbb{R}^{n_\gamma}$, where $n_\gamma = n(n-1)/2$, $\boldsymbol{\delta}_t = 2\log(\operatorname{diag}(\mathbf{D}_t))$, and $\boldsymbol{\gamma}_t = \operatorname{vecl}(\log(\mathbf{C}_t))$.[16] 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

[15] The linear operator $\operatorname{diag}(\cdot)$, 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.

[16] The linear operator $\operatorname{vecl}(\cdot)$ 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.

parameters.[17] Consider the following law of motion for $(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=2}^T$:

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\nu}_t, \text{ with } \boldsymbol{\nu}_t \sim \mathrm{N}(\mathbf{0}_{nm}, \mathbf{V}_{\boldsymbol{\beta}}) \text{ and } \boldsymbol{\beta}_t = \mathrm{vec}(\mathbf{B}_t), \tag{9}$$

$$\boldsymbol{\delta}_t = \boldsymbol{\delta}_{t-1} + \boldsymbol{\eta}_t, \text{ with } \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_n, \mathbf{V}_{\boldsymbol{\delta}}), \tag{10}$$

$$\boldsymbol{\gamma}_t = \boldsymbol{\gamma}_{t-1} + \boldsymbol{\zeta}_t, \text{ with } \boldsymbol{\zeta}_t \sim \mathrm{N}\left(\mathbf{0}_{n_\gamma}, \mathbf{V}_{\boldsymbol{\gamma}}\right), \tag{11}$$

where $\mathbf{V}_{\boldsymbol{\beta}}$ is a symmetric definite positive $nm \times nm$ matrix, $\mathbf{V}_{\boldsymbol{\delta}} = \mathrm{diag}(\mathrm{V}_{\delta,1}, \ldots, \mathrm{V}_{\delta,\mathrm{n}})$ is a diagonal definite positive $n \times n$ matrix and $\mathbf{V}_{\boldsymbol{\gamma}} = \mathrm{diag}(\mathrm{V}_{\gamma,1}, \ldots, \mathrm{V}_{\gamma,\mathrm{n}_\gamma})$ is a diagonal definite positive $n_\gamma \times n_\gamma$ matrix. In addition, we assume that $\boldsymbol{\beta}_1 \sim \mathrm{N}(\mathbf{m}_{\boldsymbol{\beta}_1}, \mathbf{V}_{\boldsymbol{\beta}_1})$, 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; $\boldsymbol{\delta}_1 \sim \mathrm{N}(\mathbf{m}_{\boldsymbol{\delta}_1}, \mathbf{V}_{\boldsymbol{\delta}_1})$, 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 $\boldsymbol{\gamma}_1 \sim \mathrm{N}(\mathbf{m}_{\boldsymbol{\gamma}_1}, \mathbf{V}_{\boldsymbol{\gamma}_1})$, 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:

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

which is a vector of dimension $n^{RC} = \frac{nm(nm+1)}{2} + n + n_\gamma + nm + \frac{nm(nm+1)}{2} + 2n + 2n_\gamma$. We denote the above prior over $(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T$ by

$$p^{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) = p^{RC}((\mathbf{B}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) p^{RC}((\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}),$$

which we call the Random Correlations prior, where $p^{RC}((\mathbf{B}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) = \prod_{t=2}^T p^{RC}(\mathbf{B}_t \mid \mathbf{B}_{t-1}, \boldsymbol{\phi}^{RC}) p^{RC}(\mathbf{B}_1 \mid \boldsymbol{\phi}^{RC})$ and

$$p^{RC}((\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) = \prod_{t=2}^T p^{RC}(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t \mid \boldsymbol{\delta}_{t-1}, \boldsymbol{\gamma}_{t-1}, \boldsymbol{\phi}^{RC}) p^{RC}(\boldsymbol{\delta}_1, \boldsymbol{\gamma}_1 \mid \boldsymbol{\phi}^{RC}) =$$
$$\left(\prod_{t=2}^T p^{RC}(\boldsymbol{\delta}_t \mid \boldsymbol{\delta}_{t-1}, \boldsymbol{\phi}^{RC}) p^{RC}(\boldsymbol{\gamma}_t \mid \boldsymbol{\gamma}_{t-1}, \boldsymbol{\phi}^{RC})\right) p^{RC}(\boldsymbol{\delta}_1 \mid \boldsymbol{\phi}^{RC}) p^{RC}(\boldsymbol{\gamma}_1 \mid \boldsymbol{\phi}^{RC}).$$

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:

$$p_R^{RC}((\mathbf{B}_t, \boldsymbol{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) = v_{g^{RC}}((\boldsymbol{\Sigma}_t)_{t=1}^T) p^{RC}((\mathbf{B}_t, g^{RC}(\boldsymbol{\Sigma}_t))_{t=1}^T \mid \boldsymbol{\phi}^{RC}), \tag{12}$$

[17] We will also use the function g^{RC} to denote a mapping from $\boldsymbol{\Sigma}_t$ to $(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)$.

where the volume element $v_{g^{RC}}((\mathbf{\Sigma}_t)_{t=1}^T)$ can be computed by numerical differentiation.[18]

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 condition 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:

$$p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) = 2^{\frac{n(n+1)T}{2}} \Big(\prod_{t=1}^{T} \mid \det(\mathbf{A}_t) \mid \Big)^{-(2n+m+1)} \frac{p_R^{RC}(\pi(f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T))) \mid \boldsymbol{\phi}^{RC})}{v(\mathcal{O}_n)^T} \tag{13}$$

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

It will be useful to introduce more notation. The mapping $f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T)$ combined with $\bar{g}^{RC}((\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T) = (\mathbf{B}_t, g^{RC}(\mathbf{\Sigma}_t), \mathbf{Q}_t)_{t=1}^T$, which we will denote $(\bar{g}^{RC} \circ f_h)((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) = (\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T$, constitutes the key ingredient of the Gibbs Sampler for inference based on time-varying SVARs identified with sign restrictions developed in this paper, see Section 5.

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 $\mathbf{\Sigma}_t = \mathbf{\Xi}_t^{-1} \mathbf{\Omega}_t \mathbf{\Omega}_t' (\mathbf{\Xi}_t')^{-1}$, where $\mathbf{\Omega}_t$ is a positive diagonal matrix and $\mathbf{\Xi}_t$ is a lower triangular matrix with ones along

[18] Because $\mathbf{\Sigma}_t$ is symmetric, there is an implicit linear restriction on $\mathbf{\Sigma}_t$. 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 et al. (2018). Either approach will give identical answers, though using Theorem 3 may be more efficient numerically.

the diagonal. The prior over $(\mathbf{B}_t, \mathbf{\Xi}_t, \mathbf{\Omega}_t)_{t=1}^T$ defined in Primiceri (2005), which we label

$$p^P((\mathbf{B}_t, \mathbf{\Xi}_t, \mathbf{\Omega}_t)_{t=1}^T \mid \boldsymbol{\phi}^P) = p^P((\mathbf{B}_t)_{t=1}^T \mid \boldsymbol{\phi}^P) p^P((\mathbf{\Xi}_t, \mathbf{\Omega}_t)_{t=1}^T \mid \boldsymbol{\phi}^P),$$

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:

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

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 condition given by Equation (3). We denote such prior by $p_S^P((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^P)$:

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

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 $n!$ different elements of the class and posteriors, where n is the number of variables.[19] At this point is also important to highlight Cogley and Sargent (2005), who present a similar approach with a time-invariant Ξ. Hence, their model could also be used as an alternative time-varying reduced-form model to be combined with the uniform prior over the sequences of orthogonal matrices to induce an alternative element of our class.

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 relays on the discounted Wishart stochastic volatility model to directly define a prior over the time-varying

[19] It is important to highlight that if one considers Primiceri's (2005) as a time-varying SVAR model, it does not belong to the order-invariant class. This is due to the recursive identification.

reduced-form parameters, which we label $p_R^{DW}((\mathbf{B}_t, \mathbf{\Sigma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{DW})$. 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:

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

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 et al. (2015) pointing out that a natural way to determine the impact of priors is to assess their implied out-of-sample forecasting performance. In a similar environment, Arias et al. (2023) shows that the Random Correlations prior outperforms most of the orderings of the variables in Primiceri (2005) and Bognanni (2018). Unfortunately, given the number of variables in the current framework we cannot compare the Random Correlations prior against all the orderings of the variables in Primiceri (2005). Finally, as explained in Arias et al. (2023), we also could 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 the procedures, we favor the Random Correlations prior approach presented here because it preserves the spirit of the random walk modeling in Primiceri (2005). Chan et al. (2021) also describes 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. Oftentimes we will refer to this distribution as the desired target distribution.

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 $\boldsymbol{S}_{S,t}(\mathbf{A}_t, \mathbf{F}_t)$ 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 $\boldsymbol{S}_{S,t}(\mathbf{A}_t, \mathbf{F}_t) > 0$. For instance, the function could be a collection of impulse responses of various variables to various shocks at various horizons. 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 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: e.g., the Federal Reserve targeted non-borrowed reserves during the early years of Chair 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.

Let $[\boldsymbol{S}_{S,t}(\mathbf{A}_t, \mathbf{F}_t) > 0]$ be an indicator function that equals 1 if the sign restrictions are satisfied at time t, and 0 otherwise. In addition, let $\boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0$ denote the collection of sign restrictions, so that $[\boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0] = \prod_{t=1}^T [\boldsymbol{S}_{S,t}(\mathbf{A}_t, \mathbf{F}_t) > 0]$, and $\mathbb{O}_T = \{(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T : [\boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0] = 1\}$. Equipped with these definitions, we can formally state that our objective is to sample from the posterior

$$p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0)$$

defined as

$$\frac{p((\mathbf{y}_t)_{t=1}^T \mid (\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T)\left[\boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0\right] p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) p(\boldsymbol{\phi}^{RC})}{\int \int_{\mathbb{O}_T} p((\mathbf{y}_t)_{t=1}^T \mid (\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) d(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T p(\boldsymbol{\phi}^{RC}) d\boldsymbol{\phi}^{RC}}. \tag{14}$$

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 $\boldsymbol{S}_{S,t}()$ 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. The idea behind the algorithm below is to sample several sequences of time-varying orthogonal reduced-form parameters and then keep only the draws that satisfy the sign restrictions.

Algorithm 1. *This algorithm draws from* $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0)$.

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* $(\mathbf{Q}_t^{i,m})_{t=1}^T$ *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* $\left[\boldsymbol{S}_S((\mathbf{A}_t^{i,m}, \mathbf{F}_t^{i,m})_{t=1}^T) > 0\right] = 1$.

Step 2 draws $((\mathbf{B}_t^i, \boldsymbol{\delta}_t^i, \boldsymbol{\gamma}_t^i)_{t=1}^T, \boldsymbol{\phi}^{i,RC})$ from $p^{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T)$ and sets $(\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i)_{t=1}^T = (\mathbf{B}_t^i, (g^{RC})^{-1}(\boldsymbol{\delta}_t^i, \boldsymbol{\gamma}_t^i))_{t=1}^T$. These draws approximate the desired distribution via a Gibbs Sampler algorithm (see Arias et al., 2023, for details). As mentioned in Section 4.1, using a different prior over $(\mathbf{B}_t^i, \boldsymbol{\Sigma}_t^i)_{t=1}^T$ 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); Cogley and Sargent (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 infeasible. 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 et al., 2020) modify Step 3. Let $\boldsymbol{S}_{OR,t}(\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t) = \boldsymbol{S}_{S,t}(f_h^{-1}(\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t))$ and $\mathbb{O}_n(\mathbf{B}_t, \boldsymbol{\Sigma}_t) = \{\mathbf{Q}_t : [\boldsymbol{S}_{OR,t}(\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t) > 0] = 1\}$. Current algorithms draw $(\mathbf{Q}_t^{i,m})_{m=1}^M$ independently from the uniform distribution over $\mathbb{O}_n(\mathbf{B}_t, \boldsymbol{\Sigma}_t)$ for $1 \leq m \leq M$ and $1 \leq t \leq T$. Given a draw from the specified posterior over the sequences of reduced-form parameters, this is implemented by drawing $\mathbf{Q}_t$ independently from the uniform distribution over $\mathcal{O}_n$ until one obtains a draw such that $\mathbf{Q}_t \in \mathbb{O}_n(\mathbf{B}_t, \boldsymbol{\Sigma}_t)$, for $1 \leq t \leq T$.[20] The current algorithms can be written as follows:

Algorithm 2. *This algorithm draws from a posterior distribution of* $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ *conditional on the sign restrictions.*

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

[20] 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.

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

3. *Draw* $\mathbf{Q}_t^i$ *uniformly from the set* $\mathbb{O}_n(\mathbf{B}_t^i, \mathbf{\Sigma}_t^i)$ *for* $1 \leq t \leq T$.

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.*

Step 3 can be done by drawing $\mathbf{Q}_t$ from $\mathcal{O}_n$ until obtaining one that satisfies $\boldsymbol{S}_{S,t}(f_h^{-1}(\mathbf{B}_t^i, \mathbf{\Sigma}_t^i, \mathbf{Q}_t)) > \mathbf{0}$ is found. The computational benefit comes from the *conditional* nature of the prior distribution for $\mathbf{Q}_t$. However, it has an 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 $[\boldsymbol{S}_{OR,t}(\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t) > 0]$:

$$p(\mathbf{Q}_t \mid \mathbf{B}_t, \mathbf{\Sigma}_t) = \frac{[\boldsymbol{S}_{OR,t}(\mathbf{B}_t, \mathbf{\Sigma}_t, \mathbf{Q}_t) > 0]}{v(\mathcal{O}_n(\mathbf{B}_t, \mathbf{\Sigma}_t))} \text{ for } 1 \leq t \leq T \tag{15}$$

which demonstrates that current algorithms define a density for the orthogonal matrices conditional on the reduced-form parameters that are not uniform. To see that, notice that $v(\mathcal{O}_n(\mathbf{B}_t, \mathbf{\Sigma}_t))$ depends on the reduced-form parameters. Thus, current algorithms use this prior:

$$\Big(\prod_{t=1}^T |\det(\mathbf{A}_t)| \Big)^{-(2n+m+1)} \frac{p_R^{RC}(\pi(f_h((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T))) \mid \boldsymbol{\phi}^{RC})}{\prod_{t=1}^T v(\mathcal{O}_n(\pi(f_h(\mathbf{A}_t, \mathbf{F}_t))))}, \tag{16}$$

instead of using $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC})$. To see this, the reader should compare Equations (13) and (16). In particular, in the case of Equation (16) the space of integration depends on the particular sequence of time-varying reduced-form parameters being drawn. Hence, the prior implied by the current algorithms does not satisfy Proposition 3. In this particular version, the posterior implied by Algorithm 2 will misrepresent $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0)$, as it will overrepresent sequences of structural parameters with low $\prod_{t=1}^T v(\mathcal{O}_n(\pi(f_h(\mathbf{A}_t, \mathbf{F}_t))))$ and it will underrepresent sequences with high $\prod_{t=1}^T v(\mathcal{O}_n(\pi(f_h(\mathbf{A}_t, \mathbf{F}_t))))$. Because our objective is to draw from $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0)$, one could use importance sampling weights proportional to $\frac{\prod_{t=1}^T v(\mathcal{O}_n(\pi(f_h(\mathbf{A}_t, \mathbf{F}_t))))}{v(\mathcal{O}_n)^T}$ to re-sample the outcome of Algorithm 2. However, as explained in Durbin and Koopman (2002), 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. Our sampling approach is based on single-move sampling along the lines of Koop and Potter (2011). The key insight is to draw from a sequence of conditional distributions that have the elliptical slice sampling of Murray et al. (2010) as a critical component. This will allow us to break the infeasibility of Algorithm 1. Our algorithm generates posterior draws based on the prior distribution that satisfies the conditions in Proposition 3. We will present the implementation details of the main steps of this algorithm, relegating the details to the Online Appendix.

Let $\boldsymbol{S}_{RC,t}(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t) = \boldsymbol{S}_{S,t}((\bar{g}^{RC} \circ f_h)^{-1}(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t))$ and let $[\boldsymbol{S}_{RC,t}(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t) > 0]$ be the indicator function that equals 1 if the sign restrictions are satisfied at time t and 0 otherwise. In addition, let $[\boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0] = \prod_{t=1}^T [\boldsymbol{S}_{RC,t}(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t) > 0]$ denote the collection of sign restrictions and let $\mathbb{O}_T^{RC} = \{((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T : [\boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0] = 1\}$ be the set of sequences that satisfy the sign restrictions.

The objective is to sample from Equation (14). By the theory in Section 3, this can be accomplished by sampling from the posterior

$$p^{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0)$$

defined as:

$$\frac{p((\mathbf{y}_t)_{t=1}^T \mid (\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T)\,[\boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0]\, p^{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) p(\boldsymbol{\phi}^{RC})}{\int \int_{\mathbb{O}_T^{RC}} p((\mathbf{y}_t)_{t=1}^T \mid (\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T) p^{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T \mid \boldsymbol{\phi}^{RC}) d(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T p(\boldsymbol{\phi}^{RC}) d\boldsymbol{\phi}^{RC}},$$

and then mapping $(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T$ to $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T$ using the mapping $(\bar{g}^{RC} \circ f_h)^{-1}$. This procedure is described by Algorithm 3.

Algorithm 3. *This algorithm draws from* $p_S^{RC}((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}^{RC} \mid (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T) > 0)$.

1. *Let* $I > 1$ *and set* $i = 1$ *and assign initial values to* $(\boldsymbol{\delta}_t^{i-1}, \boldsymbol{\gamma}_t^{i-1}, \mathbf{Q}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}$.

2. *Draw* $(\mathbf{B}_t^i)_{t=1}^T$ *from*

$$p^{RC}((\mathbf{B}_t)_{t=1}^T \mid (\boldsymbol{\delta}_t^{i-1}, \boldsymbol{\gamma}_t^{i-1}, \mathbf{Q}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}, (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t^{i-1}, \boldsymbol{\gamma}_t^{i-1}, \mathbf{Q}_t^{i-1})_{t=1}^T) > 0).$$

3. *Draw* $(\mathbf{Q}_t^i)_{t=1}^T$ *from*

$$p^{RC}((\mathbf{Q}_t)_{t=1}^T \mid (\mathbf{B}_t^i, \boldsymbol{\delta}_t^{i-1}, \boldsymbol{\gamma}_t^{i-1})_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}, (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_{RC}((\mathbf{B}_t^i, \boldsymbol{\delta}_t^{i-1}, \boldsymbol{\gamma}_t^{i-1}, \mathbf{Q}_t)_{t=1}^T) > 0).$$

4. *Draw* $(\boldsymbol{\delta}_t^i, \boldsymbol{\gamma}_t^i)_{t=1}^T$ *from*

$$p^{RC}((\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T \mid (\mathbf{B}_t^i, \mathbf{Q}_t^i)_{t=1}^T, \boldsymbol{\phi}^{i-1,RC}, (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_{RC}((\mathbf{B}_t^i, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t^i)_{t=1}^T) > 0).$$

5. *Draw* $\boldsymbol{\phi}^{i,RC}$ *from*

$$p^{RC}(\boldsymbol{\phi}^{RC} \mid (\mathbf{B}_t^i, \boldsymbol{\delta}_t^i, \boldsymbol{\gamma}_t^i, \mathbf{Q}_t^i)_{t=1}^T, (\mathbf{y}_t)_{t=1}^T).$$

6. *Set* $(\mathbf{A}_t^i, \mathbf{F}_t^i)_{t=1}^T = (\bar{g}^{RC} \circ f_h)^{-1}((\mathbf{B}_t^i, \boldsymbol{\delta}_t^i, \boldsymbol{\gamma}_t^i, \mathbf{Q}_t^i)_{t=1}^T)$.

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

Step 2 draws $\mathbf{B}_t$ from $p^{RC}(\mathbf{B}_t \mid \mathbf{B}_{-t}, (\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T, \boldsymbol{\phi}^{RC}, (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0)$ for $1 \leq t \leq T$. These densities are truncated normal and, hence, standard algorithms can be used to draw from them. Step 3 is implemented by drawing $\mathbf{Q}_t$ from $p^{RC}(\mathbf{Q}_t \mid \mathbf{Q}_{-t}, (\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)_{t=1}^T, \boldsymbol{\phi}^{RC}, (\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0)$ for $1 \leq t \leq T$. These densities are truncated uniform over the set of orthogonal matrices $\frac{[\boldsymbol{S}_{RC,t}(\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t) > 0]}{v(\mathcal{O}_n(\mathbf{B}_t, (g^{RC})^{-1}(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)))}$ for $1 \leq t \leq T$ and, hence, standard algorithms can be used to draw from them. Step 4 is implemented by drawing $(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t)$ from $p^{RC}(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t \mid \boldsymbol{\delta}_{-t}, \boldsymbol{\gamma}_{-t}, (\mathbf{B}_t, \mathbf{Q}_t)_{t=1}^T, \boldsymbol{\phi}^{RC}(\mathbf{y}_t)_{t=1}^T, \boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0)$ for $1 \leq t \leq T$. We use the elliptical slice sampler to do that. The key to the approach is the fact that $p^{RC}(\boldsymbol{\delta}_t, \boldsymbol{\gamma}_t \mid \boldsymbol{\delta}_{-t}, \boldsymbol{\gamma}_{-t}, \boldsymbol{\phi}^{RC}) = p^{RC}(\boldsymbol{\delta}_t \mid \boldsymbol{\delta}_{-t}, \boldsymbol{\phi}^{RC}) p^{RC}(\boldsymbol{\gamma}_t \mid \boldsymbol{\gamma}_{-t}, \boldsymbol{\phi}^{RC})$ is normal for $1 \leq t \leq T$. When it is not normal, one may need to use a different strategy for posterior sampling, and we illustrate how this can be approached in Section **??** of the Online Appendix. Step 5 is based on a standard hyperparameter updating. All the details are described in the Online Appendix. The fundamental insight of our Gibbs Sampler is that we need to condition on $\boldsymbol{S}_{RC}((\mathbf{B}_t, \boldsymbol{\delta}_t, \boldsymbol{\gamma}_t, \mathbf{Q}_t)_{t=1}^T) > 0$ at every step. Without such conditioning at every step, our approach will not be drawing from the correct posterior distribution. Related to this point, it is important to notice that the density used in Step 3 is the same density displayed in Equation (15). This highlights that readers should not infer that any use of such a density is always incorrect; the key issue is how the density is used in the context of the algorithm.

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. 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 (2021)? 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. Often, we will refer to the latter as the credit spread. 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. Appendix B provides a detailed description of the data sources.

Our prior over $\boldsymbol{\phi}^{RC}$ is described in Appendix C. For ease of exposition, we present a summary. We set $\mathbf{m}_{\boldsymbol{\beta}_1}$ equal to the maximum likelihood estimate of a constant parameter reduced-form VAR—featuring the same variables, constant, and lags as our time-varying model—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 4 times the unbiased estimator for the variance of $\hat{\mathbf{B}}$, as in Primiceri (2005). 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 $\boldsymbol{V}_{\boldsymbol{\gamma}_1}$. The variances $\boldsymbol{V}_{\boldsymbol{\delta}_1}$ and $\boldsymbol{V}_{\boldsymbol{\gamma}_1}$ are set equal

to 4 times the variance implied by the Delta method. Turning to the parameters governing the step sizes of the processes for $\boldsymbol{\beta}_t$, $\boldsymbol{\delta}_t$, and $\boldsymbol{\gamma}_t$ ($\mathbf{V}_{\boldsymbol{\beta}}$, $\mathbf{V}_{\boldsymbol{\delta}}$, and $\mathbf{V}_{\boldsymbol{\gamma}}$ respectively), we impose an Inverse-Wishart prior for $\mathbf{V}_{\boldsymbol{\beta}}$ and an Inverse-Gamma prior for each of the diagonal entries of $\mathbf{V}_{\boldsymbol{\delta}}$ and $\mathbf{V}_{\boldsymbol{\gamma}}$. The scale parameters for these priors are chosen to be constant fractions of the maximum likelihood estimate variances for $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, and $\boldsymbol{\gamma}$ in the constant parameter reduced-form VAR over the training sample described above. We follow Primiceri (2005) when setting the degrees of freedom (for the Inverse-Wishart prior) and the shape parameters (for the Inverse-Gamma priors); they are set so that the priors are diffuse and uninformative. Our results are based on one independent chain obtained using Algorithm 3. The chain consists of 1,000,000 draws, we keep one every 50-th draw of the structural parameters. Of the resulting 20,000 draws, we discard the first 5,000 draws. The Online Appendix shows some convergence results.

6.2 Identification

We identify the parameters of one of the RC-SVAR equations, which we label the monetary policy equation, by incorporating the insights of the identification strategies proposed by Uhlig (2005) and Arias et al. (2019), and at the same time allowing for time variation in both the systematic component of monetary policy and the variance of the monetary policy shock. This is critical because the Federal Reserve's 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., 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 et al. (2019).[21] 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.

Without loss of generality, 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

$$\mathrm{r}_t = \underbrace{\psi_{\Delta y,t}\Delta y_t + \psi_{\pi,t}\pi_t + \psi_{\Delta m,t}\Delta m_t + \psi_{\mathrm{cs},t}\mathrm{cs}_t}_{\text{Systematic Component}} + \underbrace{\sigma_{r,t}\varepsilon_{r,t}}_{\text{Shock}}, \tag{17}$$

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, cs_t is the corporate credit spread, $\varepsilon_{r,t}$ is the monetary policy shock, $\psi_{\Delta y,t} = -a_{t,11}a_{t,31}^{-1}$, $\psi_{\pi,t} = -a_{t,21}a_{t,31}^{-1}$, $\psi_{\Delta m,t} = -a_{t,41}a_{t,31}^{-1}$, $\psi_{\mathrm{cs},t} = -a_{t,51}a_{t,31}^{-1}$, and $\sigma_{r,t} = a_{t,31}^{-1}$, with $a_{t,ij}$ denoting the (i,j) entry of $\mathbf{A}_t$. Sometimes we will refer to the coefficients $(\psi_{\Delta y,t}, \psi_{\pi,t}, \psi_{\Delta m,t}, \psi_{\mathrm{cs},t})$ as contemporaneous elasticities. Importantly, this equation clarifies that the monetary policy shock represents a deviation from a policy rule. Equation (17) is enough to describe the contemporaneous elasticities at the center stage of the monetary SVAR literature (e.g., Bernanke and Mihov, 1998). However, there are cases in which the interest is in using the monetary policy equation to compute the long-run response of the federal funds rate to a permanent increase in inflation. When computing these long-run responses, we will also need to consider $\psi_{r,t}^{(1)} = f_{t,41}a_{t,31}^{-1}$, $\psi_{r,t}^{(2)} = f_{t,91}a_{t,31}^{-1}$, $\psi_{\pi,t}^{(1)} = f_{t,31}a_{t,31}^{-1}$ and $\psi_{\pi,t}^{(2)} = f_{t,81}a_{t,31}^{-1}$, with $f_{t,ij}$ denoting the (i,j) entry of $\mathbf{F}_t$ and $\psi_{\pi,t}^{(1)}$ and $\psi_{\pi,t}^{(2)}$ denoting the effect of π_{t-1} and π_{t-2} on r_t, respectively, and $\psi_{r,t}^{(1)}$ and $\psi_{r,t}^{(2)}$ denoting the effect of r_{t-1} and r_{t-2} on r_t, respectively.

Following Primiceri (2005), we define $\Psi_{r,t}^{(h)}$, with $h > 0$, to be the cumulative response of the fed funds rate in period $h+t$ after a permanent 1 percentage point increase in inflation in

[21] 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.

period t. More specifically, we define:

$$\Psi_{r,t}^{(h)} = \psi_{r,t}^{(1)}\Psi_{r,t}^{(h-1)} + \psi_{r,t}^{(2)}\Psi_{r,t}^{(h-2)} + \psi_{\pi,t}^{(2)} + \psi_{\pi,t}^{(1)} + \psi_{\pi,t} \text{ for } h > 2,$$

where $\Psi_{r,t}^{(2)} = \psi_{r,t}^{(1)}\Psi_{r,t}^{(1)} + \psi_{\pi,t}^{(1)} + \psi_{\pi,t}$ and $\Psi_{r,t}^{(1)} = \psi_{\pi,t}$ and we approximate the long-run response of the fed funds rate to inflation by $\Psi_{r,t}^{(60)}$. Let us now 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.[22] 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).[23]

Restriction 2. *Consider the following restrictions on the contemporaneous coefficients of Equation (17), $\psi_{\Delta y,t} \in (0,4)$, $\psi_{\pi,t} \in (0,4)$, $\psi_{\Delta m,t} \in (0,4)$, and $\psi_{cs,t} \in (-4,0)$. In addition, we restrict the long-run response of the fed funds rate to inflation to be positive and respect the Taylor principle since the mid-1980s.*

Restriction 2, inspired by Arias et al. (2019), addresses these concerns by restricting the contemporaneous rule coefficients as well as the long-run response of the federal funds rate to

[22]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.

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

inflation. The signs and bounds on the contemporaneous reaction of the federal funds rate to output growth and inflation follow directly from Arias et al. (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 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 be 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).

In addition to restricting the short-run elasticities, we restrict the long-run response of the fed funds rate to inflation to be positive and to respect the Taylor principle since 1984Q1. The former is inspired in the New Keynesian framework. The rationale for the Taylor principle restriction is inspired in Clarida et al. (2000); Lubik and Schorfheide (2004), pointing to a shift in the long-run response of the policy rate to inflation in the early 1980s.

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) provides 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 et al. (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 target 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 (17) reemerged only after the federal funds rate lift-off was 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 federal 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 output growth 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 0.01 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 also find high elasticities in the early 1970s under the first years of Burns's tenure and around the 2000s under Greenspan. The former 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, 2024) 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. The high elasticities around 1999-2000 are consistent concerns about inflationary pressures mentioned in the FOMC statement of the time. 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. For example, these responses were as low as 0.14 during Janet Yellen's leadership.

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 1980s were concerned with the reliability of monetary aggregates. For example, Lindsey et al. (2013) emphasize the following quote from Chair 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]".

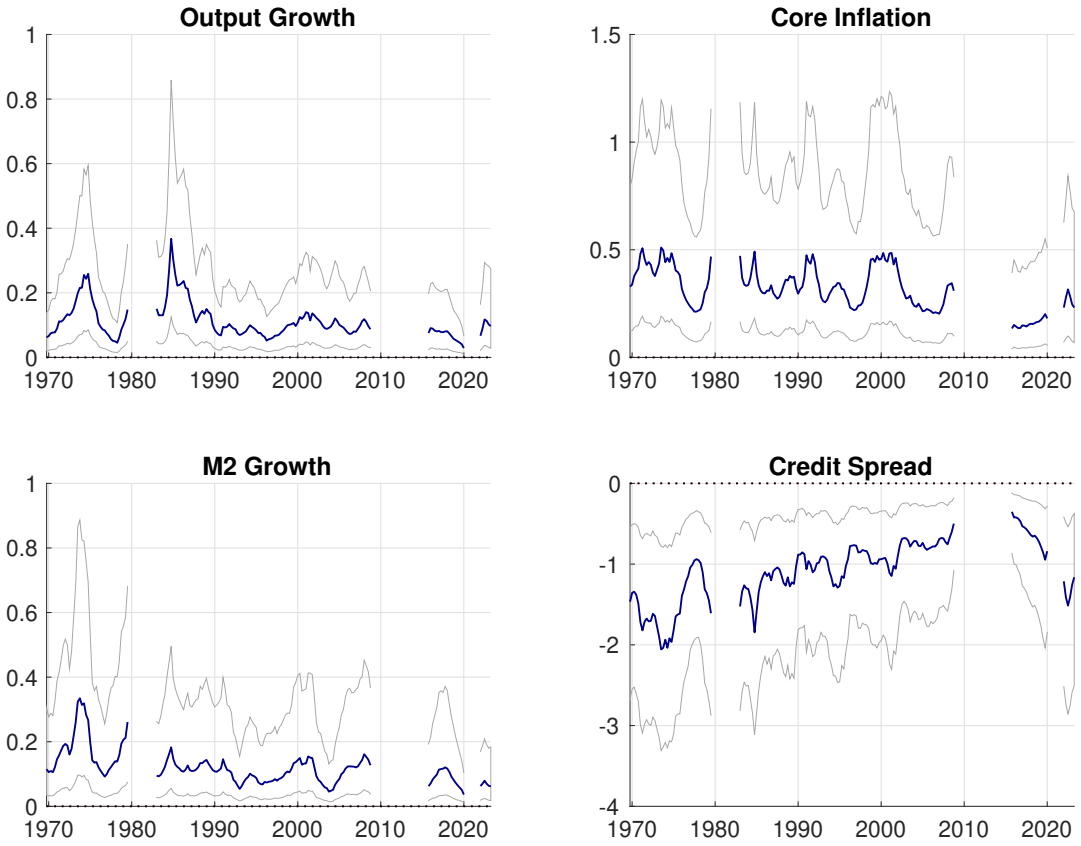

Figure 1: Contemporaneous Elasticities

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 −1.3 at the beginning of Burns's tenure to −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 in responsiveness ended after the financial crisis. By mid-2023, the response was about -1.2, which aligns with the posterior median estimates for the 1990s and early 2000s. While Caldara and Herbst (2019) focuses 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.

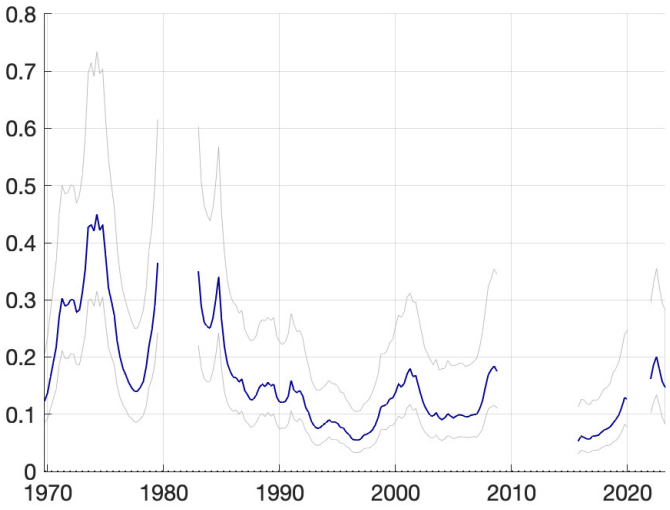

Figure 2: Standard Deviation of the Monetary Policy Shock

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.

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, 2023). 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 federal funds rate, output growth, and inflation over the period 2022Q2-2023Q2, using data from 2022Q1. In addition, the figure shows the cumulative contribution of the structural shocks to the forecast error in these projections. The forecast and the cumulative shock contributions are constructed using point-wise means

conditional on the distribution of structural parameters corresponding to 2022Q1. The former are represented by dotted lines and the latter are represented by the colored bars. We depict the contribution of monetary policy shocks using red bars, while the contributions from non-monetary policy shocks are shown in yellow bars. For each quarter between 2022Q2 and 2023Q2, the gap between the data (shown by solid lines) and the forecast is equal to the sum of the yellow and red bars.

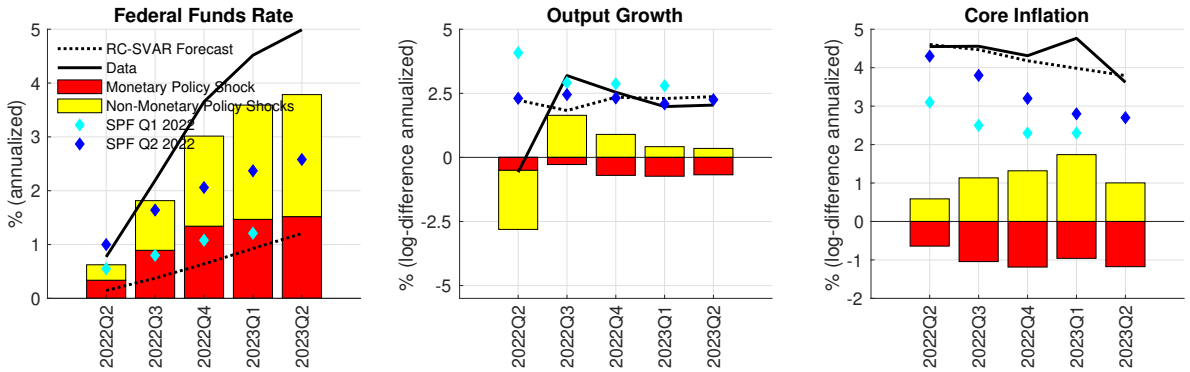

Figure 3: Monetary Policy versus Non-Monetary Policy Shocks

The forecast for the federal funds rate shows that the model significantly under-predicted its trajectory. In contrast, the projection for output growth and inflation was remarkably accurate. We contrast our model with the median predictions from the 2022Q1 and 2022Q2 Survey of Professional Forecasters (SPF).[24] 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 is broadly in line with the SPF for the case of the federal funds rates and output growth, while the SPF does worse than the model when forecasting 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

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

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. Even so, monetary policy shocks also played a role, amounting to about 150 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 Figure

	2022Q2	2022Q3	2022Q4	2023Q1	2023Q2
Predictable	0.15	0.37	0.64	0.93	1.20
Unpredictable due to Systematic	0.29	0.92	1.67	2.12	2.27
Unpredictable due to MP Shocks	0.34	0.89	1.34	1.47	1.52
Federal Funds Rate	0.77	2.19	3.65	4.52	4.99

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

3, non-monetary policy shocks are also the main contributor to the unexpected changes in output growth and inflation: since the second half of 2022 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.

Our analysis restricts monetary shocks to be one-off deviations from the estimated monetary policy equation; intuitively, however, such deviations could be either transitory or persistent, and differentially so at different points in time. Hence, it is natural to ask how our conclusions on the importance of systematic monetary policy versus monetary shocks in explaining nominal interest rate movements post-2021 are affected by allowing for multiple distinct monetary policy shocks. To assess this, the Online Appendix re-visits our analysis by disentangling exogenous variation at the short end and long end of the yield curve. We find that the share of the unexpected changes in the federal funds rate that can be attributed to non-monetary policy shocks is in line with the results presented in this section. In addition, the monetary policy shock in the baseline is deconstructed in similar proportions into the short end and long-term monetary policy shocks: the sum of both short- and long end monetary policy shocks seems to be equally important contributors to the unexpected variations in the

variables under analysis than the specification featuring the term spread.

6.5 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 (2006a) 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 et al. (2024); these address the expectational concerns of the Lucas critique.

The results obtained under both types of counterfactual simulations are similar. We will focus on the first type and discuss the results obtained with the second type in the Online Appendix. More specifically, we 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 4 shows the point-wise medians and the 68 percent point-wise probability bands for each counterfactual. 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

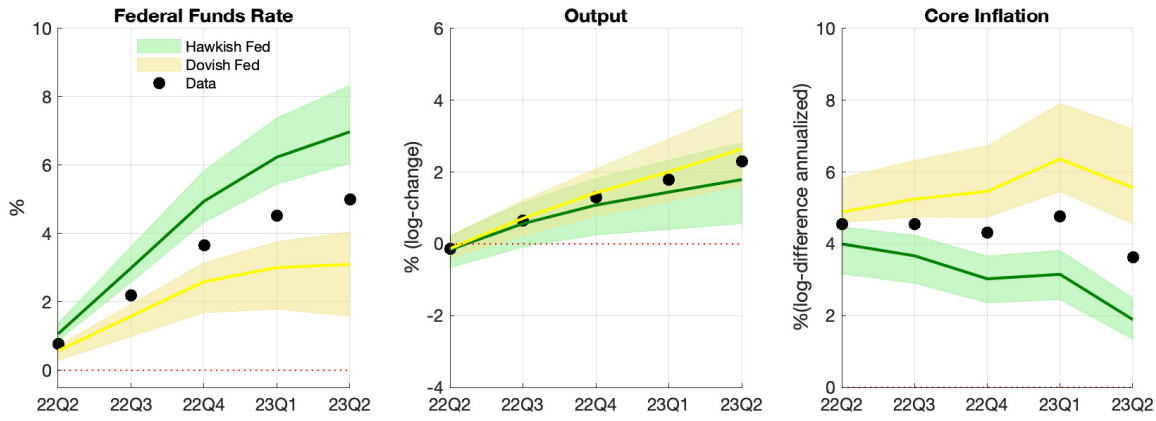

Figure 4: Counterfactuals

Fed counterfactual, inflation would have quickly decreased in terms of economic activity at a relatively small cost. In the Hawkish counterfactual, the output in the second quarter of 2023 would have been about 0.5 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 1.7 percent.

Let us highlight that variations in the slope of the Phillips curve could influence the transmission mechanisms of monetary policy, potentially altering the trade-offs between output and inflation. Our counterfactual results are coherent with price responses being much larger at the end of the sample than in the mid-seventies, suggesting Chair Burns faced a more adverse trade-off than Chair Powell.[25] In addition, the Online Appendix shows that the output loss associated with the Hawkish counterfactual is somewhat larger when we use a monetary policy equation centered around the short end of the yield curve.

6.6 Robustness to Prior

The results in Section 3 demonstrate that to operate within the class of rotation-invariant time-varying SVARs models described in this paper, a researcher must use a uniform prior over $(\mathbf{Q}_t)_{t=1}^T$ with respect to the volume measure over $\mathcal{O}_n^T$. However, as emphasized by Baumeister and Hamilton (2015); Giacomini and Kitagawa (2021), it is crucial to recognize that sign restrictions yield identified sets, and the outcomes are influenced not only by these

[25]This is shown in the Online Appendix where we compare the impulse responses of output and prices in 1975Q1 and 2022Q1. We thank a referee for bringing this point to our attention.

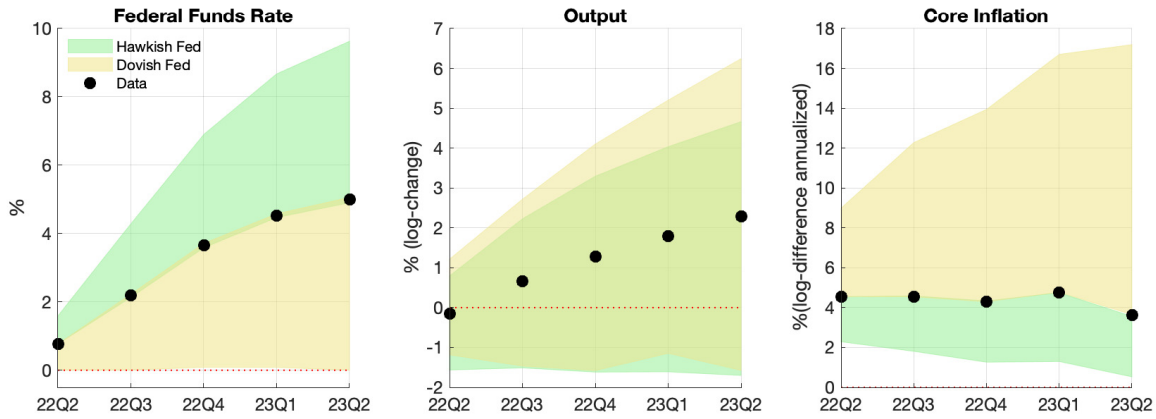

Figure 5: Identified Sets for Counterfactuals

restrictions but also by the prior imposed on the sequence of orthogonal rotation matrices.

To address this, we assess the sensitivity of the reported policy counterfactual simulations to the choice of the uniform prior. Specifically, we examine the identified sets associated with Figure 4 at the value of the point-wise posterior median sequence of reduced-form parameters.[26] The shadowed areas in Figure 5 represent the point-wise identified set. The figure shows two important things: (1) the identified sets are much wider than the probability bands and (2) the main finding of the paper is robust to reporting identified sets.

The identified sets can naturally be wider than our reported probability bands, as they represent the domain of the counterfactual outcomes in the absence of reduced-form parameter uncertainty. Although with a larger uncertainty, the identified sets still tell us that under the Dovish Fed counterfactual, the economy would have marginally overheated and inflation would have been much higher while, under the Hawkish Fed counterfactual, inflation would have been remarkably lower with a set of likely outcomes for output broadly similar to those under the Dovish counterfactual. Importantly, this exercise accounts only for identification uncertainty and not for reduced-form estimation uncertainty. Unfortunately, given the dimensionality of our problem, accounting for both is computationally infeasible.

[26]That is, we report the point-wise range of counterfactual outcomes that are observationally equivalent to the point-wise medians presented in Figure 4. Conditional on the point-wise posterior median sequence of reduced-form parameters, we simulate 50,000 sequences of orthogonal matrices that satisfy the sign restrictions. We eliminate any draws that imply negative interest rates. These represent the 0.6 percent of the total number of draws.

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. Figure 6 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.

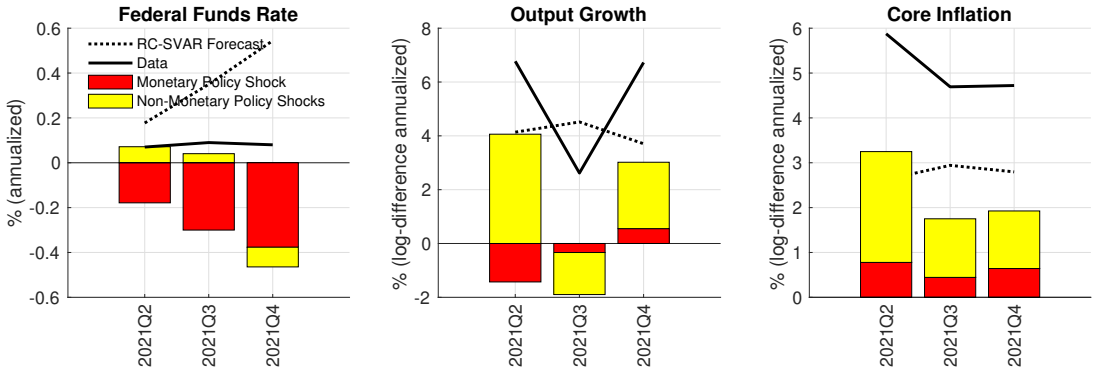

Figure 6: Monetary Policy versus Non-Monetary Policy Shocks

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 nearly 0.6 percent by the end of 2021. This projection was associated with a prediction for output growth of about 4 percent from 2021Q1 until 2021Q4 and a prediction for inflation of about 3 percent throughout the same 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 that the FOMC 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 by the end of the forecast horizon (2021Q4) is relatively small.

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 SVARs 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 et al. (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.

8 Data Availability

The data and code underlying this research is available on Zenodo at https://doi.org/10.5281/zenodo.15102165.

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.

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 J. D. Hamilton (2015). Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information. *Econometrica 83*(5), 1963–1999.

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.

Bernanke, B. S. and I. Mihov (1998). Measuring Monetary Policy. *Quarterly Journal of Economics 113*(3), 869–902.

Bognanni, M. (2018). A Class of Time-Varying Parameter Structural VARs for Inference Under Exact or Set Identification. Working Paper 18-11, Federal Reserve Bank of Cleveland.

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 (2024). Evaluating Policy Counterfactuals: A VAR-Plus Approach. Working Paper 32988, National Bureau of Economic Research.

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

Chappell, H. W. J., 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.

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

Cogley, T. and T. J. Sargent (2005, April). Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics 8*(2), 262–302.

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. (2024). Estimating the Effects of Political Pressure on the Fed: A Narrative Approach with New Data. Working Paper 32461, National Bureau of Economic Research.

Durbin, J. and S. J. Koopman (2002). A Simple and Efficient Simulation Smoother for State Space Time Series Analysis. *Biometrika 89*(3), 603–615.

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.

Giacomini, R. and T. Kitagawa (2021). Robust Bayesian Inference for Set-identified Models. *Econometrica 89*(4), 1519–1556.

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. Working Paper 2022-34, FEDS Working Paper Series.

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.

Koop, G. and S. M. Potter (2011). Time Varying VARs with Inequality Restrictions. *Journal of Economic Dynamics and Control 35*(7), 1126–1138.

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.

Lubik, T. A. and F. Schorfheide (2004, March). Testing for Indeterminacy: An Application to U.S. monetary policy. *American Economic Review 94*(1), 190–217.

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.

Murray, I., R. Adams, and D. Mackay (2010). Elliptical Slice Sampling. *Journal of Machine Learning Research: WandCP 9*, 541–548.

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.

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.

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

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. 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. (2021). 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. (1997). Bayesian Vector Autoregressions with Stochastic Volatility. *Econometrica: Journal of the Econometric Society 65*(1), 59–73.

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.

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 $p((\mathbf{y}_t)_{t=1}^T \mid (\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T, \boldsymbol{\phi}) = \prod_{t=1}^T p(\mathbf{y}_t \mid \mathbf{x}_t, \mathbf{A}_t, \mathbf{F}_t)$, 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 $(\mathbf{A}_t \mathbf{A}_t')^{-1}$. 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 $\mathbf{Q}_t \equiv \mathbf{A}_t^{-1} \tilde{\mathbf{A}}_t \in \mathcal{O}_n$, or equivalently, $\tilde{\mathbf{A}}_t = \mathbf{A}_t \mathbf{Q}_t$, for

$\mathbf{Q}_t \in \mathcal{O}_n$ and $1 \leq t \leq T$. 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 $\mathbf{x}_t$. Because, the support of the initial conditions, $(\mathbf{y}_0, \cdots, \mathbf{y}_{1-p})$, is full, the span of the support of $\mathbf{x}_1$ is all of $\mathbb{R}^m$. Thus, the span of the support of $\mathbf{x}_t$ is all of $\mathbb{R}^m$, for $1 \leq t \leq T$. Thus, $\tilde{\mathbf{F}}_t = \mathbf{F}_t \mathbf{Q}_t$ for $1 \leq t \leq T$. □

Proof of Proposition 2. To prove the proposition, we will show that the priors over the time-varying structural parameters implied by models with heteroskedastic structural shocks do not satisfy Equation (3). Implicit in Equation (3) is the fact that if $p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi}) > 0$, then $p_S((\mathbf{A}_t \mathbf{Q}_t, \mathbf{F}_t \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi}) > 0$, for every $(\mathbf{Q}_t)_{t=1}^T \in \mathcal{O}_n^T$. For models with heteroskedastic structural shocks, the prior $p_S((\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T \mid \boldsymbol{\phi})$ 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 $(\mathbf{A}_t, \mathbf{F}_t)_{t=1}^T = (\mathbf{A}\,\boldsymbol{\Psi}_t^{-\frac{1}{2}}, \mathbf{F}\,\boldsymbol{\Psi}_t^{-\frac{1}{2}})_{t=1}^T$. In particular, if the prior is positive then it must be the case that $\mathbf{A}_1^{-1}\mathbf{A}_t = \boldsymbol{\Psi}_1^{\frac{1}{2}}\boldsymbol{\Psi}_t^{-\frac{1}{2}}$ is diagonal for $1 \leq t \leq T$. Suppose that $p_S((\mathbf{A}\,\boldsymbol{\Psi}_t^{-\frac{1}{2}}, \mathbf{F}\,\boldsymbol{\Psi}_t^{-\frac{1}{2}})_{t=1}^T) > 0$. Let $(\mathbf{Q}_t)_{t=1}^T \in \mathcal{O}_n^T$ be any sequence such that $\mathbf{Q}_1 \in \mathcal{O}_n$ is not diagonal and $\mathbf{Q}_t = \mathbf{I}_n$, for $2 \leq t \leq T$. It cannot be the case that $p_S((\mathbf{A}\,\boldsymbol{\Psi}_t^{-\frac{1}{2}}\mathbf{Q}_t, \mathbf{F}\,\boldsymbol{\Psi}_t^{-\frac{1}{2}}\mathbf{Q}_t)_{t=1}^T) > 0$ because, for $2 \leq t \leq T$, $(\mathbf{A}\,\boldsymbol{\Psi}_1^{-\frac{1}{2}}\mathbf{Q}_1)^{-1}(\mathbf{A}\,\boldsymbol{\Psi}_t^{-\frac{1}{2}}\mathbf{Q}_t) = \mathbf{Q}_1'\boldsymbol{\Psi}_1^{\frac{1}{2}}\boldsymbol{\Psi}_t^{-\frac{1}{2}}$, which is not diagonal. □

Proof of Proposition 3. 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 $p_R((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi}) = p_R((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t\mathbf{P}_t)_{t=1}^T \mid \boldsymbol{\phi})$, for every $(\mathbf{P}_t)_{t=1}^T \in \mathcal{O}_n^T$. The result follows because the Equation is satisfied if and only if $p_R((\mathbf{B}_t, \boldsymbol{\Sigma}_t, \mathbf{Q}_t)_{t=1}^T \mid \boldsymbol{\phi})$ does not depend on $(\mathbf{Q}_t)_{t=1}^T$. □

B 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.

C 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 C.1 summarizes the parameters and hyperparameters of the RC-SVAR. Next, we turn to the details.

Table C.1: 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: $\boldsymbol{\psi}^{RC}$	
$\bar{\nu}_{\boldsymbol{\beta}}$	Degrees of freedom of the prior for $\mathbf{V}_{\boldsymbol{\beta}}$.
$\bar{k}_{\boldsymbol{\beta}}$	Scaling factor for the scale matrix of the prior for $\mathbf{V}_{\boldsymbol{\beta}}$.
$\bar{\nu}_{\delta}$	Shape parameter of the prior for $V_{\delta,i}$ for $i = 1, \ldots, n$.
$\bar{k}_{\delta}$	Scaling factor for scale parameter of the prior for $V_{\delta,i}$ for $i = 1, \ldots, n$.
$\bar{\nu}_{\gamma}$	Shape parameter of the prior for $V_{\gamma,i}$ for $i = 1, \ldots, n_\gamma$.
$\bar{k}_{\gamma}$	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 constant parameter 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 4 times the unbiased estimator for the variance of $\hat{\mathbf{B}}$, as in Primiceri (2005). 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 constant parameter 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),

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

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 $\boldsymbol{V}_{g(\text{vech}(\hat{\boldsymbol{\Sigma}}))}(\boldsymbol{\Sigma}) = \frac{\mathbf{D}_g(\boldsymbol{\Sigma})2\,\mathbf{D}_n^{+}(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\,\mathbf{D}_n^{+\prime}\,\mathbf{D}_g(\boldsymbol{\Sigma})'}{T}$. Thus, $\boldsymbol{m}_{\boldsymbol{\delta}_1} = 2\log\left(\text{diag}(\hat{\mathbf{D}})\right)$, $\boldsymbol{m}_{\boldsymbol{\gamma}_1} = \text{vecl}(\log \hat{\mathbf{C}})$

$$\boldsymbol{V}_{\boldsymbol{\delta}_1} = 4\begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n,n_\gamma} \end{bmatrix} \boldsymbol{V}_{g(\text{vech}(\hat{\boldsymbol{\Sigma}}))}(\hat{\boldsymbol{\Sigma}}) \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0}_{n_\gamma,n} \end{bmatrix}, \text{ and } \boldsymbol{V}_{\boldsymbol{\gamma}_1} = 4\begin{bmatrix} \mathbf{0}_{n_\gamma,n} & \mathbf{I}_{n_\gamma} \end{bmatrix} \boldsymbol{V}_{g(\text{vech}(\hat{\boldsymbol{\Sigma}}))}(\hat{\boldsymbol{\Sigma}}) \begin{bmatrix} \mathbf{I}_{n_\gamma} \\ \mathbf{0}_{n,n_\gamma} \end{bmatrix}$$

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: $\mathbf{V}_{\boldsymbol{\beta}} \sim \text{IW}(\bar{\nu}_{\boldsymbol{\beta}}\bar{k}_{\boldsymbol{\beta}}^2\boldsymbol{V}_{\boldsymbol{\beta}_1}, \bar{\nu}_{\boldsymbol{\beta}})$, $V_{\delta,i} \sim \text{IG}\left(\frac{\bar{\nu}_\delta}{2}, \frac{\bar{\nu}_\delta\bar{k}_\delta^2(e_{i,n}'\boldsymbol{V}_{\boldsymbol{\delta}_1}e_{i,n})}{2}\right)$ for $i = 1, \ldots, n$, and $V_{\gamma,i} \sim \text{IG}\left(\frac{\bar{\nu}_\gamma}{2}, \frac{\bar{\nu}_\gamma\bar{k}_\gamma^2(e_{i,n_\gamma}'\boldsymbol{V}_{\boldsymbol{\gamma}_1}e_{i,n_\gamma})}{2}\right)$ for $i = 1, \ldots, n_\gamma$ where $e_{i,x}$ denotes the i-th column of an identity matrix of dimension x. The scale parameters are chosen to be constant fractions of the maximum likelihood estimate variances of the corresponding subsample. In particular, $\bar{k}_{\boldsymbol{\beta}} = 0.01$ and $\bar{k}_\delta = \bar{k}_\gamma = 0.1$. We choose the fractions to be larger for $V_{\delta,i}$ and $V_{\gamma,i}$. This choice reflects our prior belief that most of the time variation is on the variance of the reduced-form innovations (see, for example Sims and Zha, 2006b). The shape parameters are chosen to be the smallest natural number such that

the densities are defined. The degrees of freedom, $\bar{\nu}_{\boldsymbol{\beta}}$, are set equal to 105 so that the degrees of freedom per parameter in our model are equivalent to the ones in Primiceri (2005). As a consequence, the priors are as diffuse and uninformative as possible (see Primiceri, 2005, for a similar motivation).