A New Approach to Modeling Time-Varying VARs: Applications to Forecasting and Business Cycle Dynamics

by

Chang-Jin Kim [1]
University of Washington

Frbruary, 2025

Abstract

The log-squared errors in a reduced-form time-varying-parameter VAR (TVP-VAR) model are correlated, conditional on stochastic volatilities. These correlations pose a difficulty in directly sampling the log stochastic volatilities from the full joint conditional distribution. We overcome this difficulty by developing an MCMC algorithm that builds on a multivariate extension of the mixture-of-normals approach by Kim, Shephard, and Chib (1998). In our forecasting application, the density forecasting performance of our benchmark ordering-invariant model is mostly comparable to the best performance of Primiceri's (2005) model across orderings in a four-variable system, while the point forecasts match its median performance. In our application to business cycle dynamics, we find evidence of the increasing relative importance of the permanent shock in the U.S. economy since the mid-1980s. This can be explained by the hysteresis effect, i.e., the aggregate demand shock having permanent effects, and the decreasing response of real output to the transitory part of the aggregate demand shock.

Key Words: Time-Varying VAR model, Ordering-Invariant Stochastic Volatilities, Multivariate Mixture of Normals, Hysteresis Effect.

JEL Codes: C11, C51, E32

[1] Department of Economics, University of Washington, Seattle, WA 98195, USA [Email: changjin@uw.edu].

1. Introduction

Let us consider the following representation of the structural time-varying-parameter VAR (TVP-VAR) model with stochastic volatilities, as introduced by Cogley and Sargent (2001, 2005) and Primiceri (2005): [2]

$$y_t = \Phi_{0,t} + \Phi_{1,t} y_{t-1} + \ldots + \Phi_{p,t} y_{t-p} + A_t^{-1} \Sigma_t^{*\frac{1}{2}} \varepsilon_t, \quad \varepsilon_t \sim i.i.d.N(0, I_n), \tag{1}$$

where y_t is an $n \times 1$ vector; ε_t is a vector of standardized structural shocks; $\Sigma_t^{*\frac{1}{2}}$ is a diagonal matrix with the stochastic volatilities of the structural shocks on the main diagonal; A_t is constrained to be an $n \times n$ lower-triangular matrix with units on the main diagonal; and both VAR coefficients and the log volatilities evolve according to random walk processes. For estimation, priors are elicited for the structural parameters A_t and $\Sigma_t^{*\frac{1}{2}}$, as well as for the reduced-form parameters $\Phi_{i,t}$, $i = 0, 1, ..., p$.

Since its introduction, the above model has been employed as an estimation tool for making structural inferences on various macroeconomic issues, including, but not limited to: the sources of the Great Moderation (Benati, 2008; Gali and Gambetti, 2009); the time-varying contribution of financial shocks to output growth fluctuations (Pietro, Eickmeier, and Marcellino, 2016); the changing impact of loan supply shocks on the business cycle (Gambetti and Musso, 2017); the dynamic term structure of interest rates (Fischer et al., 2023); and the effects of taxation shocks (Mumtaz and Petrova, 2023). Related models have also been applied to forecast macro variables, and researchers have demonstrated that these models often provide better forecasts compared to their homoscedastic or constant-coefficient counterparts. [3] Recently, researchers such as Carriero, Clark, and Marcellino (2019), Chan (2023), and Zheng, Ye, and Hong (2023) have extended these models to accommodate large data sets for forecasting purposes.

However, it is well known that the estimation of the above model is sensitive to the ordering of the variables in y_t due to the lower-triangularity constraint imposed on the A_t

[2] Cogley and Sargent (2001, 2005) consider a model in which A_t is replaced by a time-invariant matrix A.

[3] For applications of related models for forecasting, refer to D'Agostino, Gambetti, and Giannone (2013), Koop and Korobilis (2013), Clark and Ravazzolo (2015), and Chan and Eisenstat (2018).

matrix. Arias, Rubio-Ramírez, and Shin (2023) show that the out-of-sample forecasting performance of the model in equation (1) depends on the ordering of the variables, with the problem being particularly serious for density and interval forecasts. Chan, Koop, and Yu (2023) provide in-depth discussions on this issue, showing that the problem is likely to become more serious in large VARs. [4] Wu and Koop (2023) treat the ordering as unknown and allow the data to select the optimal ordering.

Recently, Arias et al. (2024) developed a theory of identification and algorithms for Bayesian inference in set-identified structural TVP-VAR models through sign restrictions, building on the class of models defined by Bognanni (2018). They achieve their goal by writing sequences of structural parameters as sequences of orthogonal reduced-form parameters and defining an invertible mapping between them. [5] A prerequisite to their approach is an appropriate estimation of the reduced-form TVP-VAR model, especially in the presence of stochastic volatilities. Thus, whether the goal is to make structural inferences or macroeconomic forecasts, the literature calls for an appropriate procedure for estimating a reduced-form TVP-VAR model that is invariant to the ordering of the variables.

In this paper, we present a reduced-form TVP-VAR model with ordering-invariant multivariate stochastic volatilities. We also present an MCMC procedure that is simple and straightforward to implement. Unlike the log-squared structural errors in equation (1), the log-squared reduced-form errors in a reduced-form TVP-VAR model are correlated, conditional on the stochastic volatilities. This poses a difficulty in a multivariate extension of Kim, Shephard, and Chib's (1998) univariate stochastic volatility model, which adopts the multi-move sampler based on the mixture of normals approximation to the log-$\chi^2(1)$ distribution.

We demonstrate that this difficulty can be easily resolved by leveraging the one-to-one

[4] They also present an estimation procedure for a constant-coefficients VAR model with stochastic volatilities, in which the A_t matrix in equation (1) is replaced with a time-invariant matrix A without the lower-triangularity assumption. They leverage the identification of the A matrix in the presence of stochastic volatilities, as demonstrated by Bertsche and Braun (2022).

[5] They note that the priors over time-varying reduced-form parameters and the sequences of orthogonal matrices are not independent in the existing methods, generating draws of structural parameters that do not belong to the class of rotation-invariant time-varying structural VAR models.

correspondence between the correlation matrix of the reduced-form errors and the correlation matrix of their log squares, conditional on the stochastic volatilities. This involves deriving the full conditional posterior distribution of the mixture indicator variables, which are used to approximate the joint conditional distribution of the correlated log-squared reduced-form errors with multivariate mixtures of normals. All we need for the estimation of our model is a simple and straightforward Gibbs sampling approach.

We note that we are not the first to sample stochastic volatilities from reduced-form errors to achieve ordering invariance in a model with multivariate stochastic volatilities. For example, Philipov and Glickman (2006), Golosnoy, Gribisch, and Liesenfeld (2012), and Shin and Zhong (2020) impose a prior directly on the time-varying covariance matrix (Q_t) of the reduced-form errors, assuming it evolves according to a function of a Wishart process. However, Arias, Rubio-Ramírez, and Shin (2023) show that incorporating this approach into a TVP-VAR model results in out-of-sample forecasting performance worse than Primiceri's (2005) model across all orderings. Alternatively, Asai and McAleer (2009), Arias, Rubio-Ramírez, and Shin (2023), and references therein decompose the covariance matrix into $Q_t = \Sigma_t^{1/2} R_t \Sigma_t^{1/2}$, where $\Sigma_t^{1/2}$ is a diagonal matrix and R_t is the time-varying correlation matrix. They then focus on developing algorithms for an appropriate sampling of the R_t matrix. [6] However, these approaches typically require computationally intensive algorithms for estimation. Additionally, when sampling the log stochastic volatilities, these models sometimes overlook the correlations that exist among log-squared errors, conditional on the log stochastic volatilities. Filling these gaps in the literature is an important contribution of this paper.

In our forecasting application, we demonstrate that the density forecasting performance of our benchmark ordering-invariant model, which features a constant correlation matrix and random-walk log stochastic volatilities, is mostly comparable to the best performance of Primiceri's (2005) model across 24 different orderings in a four-variable system. Our model's

[6] In particular, Arias, Rubio-Ramírez, and Shin (2023) model the time-varying correlation matrix using random-walk processes based on Archakov and Hansen's (2021) numerically invertible mapping from the space of non-singular $n \times n$ correlation matrices to an $\frac{n(n-1)}{2} \times 1$ real vector. They demonstrate that their model has an out-of-sample forecasting performance comparable to the median outcomes of Primiceri's model (2005) across orderings.

point forecasting performance is comparable to the median performance of Primiceri's (2005) model. We use the same data set and procedure as Arias, Rubio-Ramírez, and Shin (2023), who evaluated the sensitivity of the out-of-sample forecasting performance of Primiceri's (2005) model across different orderings.

We also apply our model to investigate the relative importance of permanent and transitory shocks in U.S. business cycle dynamics. Since the seminal work of Nelson and Plosser (1982), this has been a longstanding but important issue in empirical macroeconomics. However, in the existing literature that focuses on specific sample periods, the issue remains unresolved, as surveyed by Cerra, Fatás, and Saxena (2023). Our results, based on a TVP-VAR model of real output growth and the unemployment rate, show that the relative importance of the permanent shock has been increasing since the mid-1980s. We demonstrate that this can be explained by the hysteresis effect, i.e., the aggregate demand shock having permanent effects, and the decreasing response of real output to the transitory part of the aggregate demand shock. The significance of our finding is that it is based on post-war U.S. time series data, while Cerra, Fatás, and Saxena (2023) and references therein provide empirical evidence of the hysteresis effect through cross-country analyses.

The rest of the paper is organized as follows. Section 2 presents our model specification. Section 3 describes the MCMC procedure for the estimation of the model. Section 4 evaluates the forecasting performance of the model. In Section 5, we apply the proposed model and algorithm to investigate the changing business cycle dynamics in the U.S. economy. Section 6 provides concluding remarks.

2. Model Specification

We first consider the following benchmark reduced-form TVP-VAR model with random-walk log stochastic volatilities and constant correlation matrix:

$$\begin{aligned} y_t &= (I_n \otimes x_t')\beta_t + u_t, \quad u_t \sim N(0, \Sigma_t^{\frac{1}{2}} R \Sigma_t^{\frac{1}{2}}), \\ \beta_t &= \beta_{t-1} + v_t, \quad v_t \sim i.i.d.N(0, \Omega_v), \\ ln(\sigma_t) &= ln(\sigma_{t-1}) + \eta_t, \ \eta_t \sim i.i.d.N(0, \Omega_\eta), \end{aligned} \tag{2}$$

where $x_t = [1 \quad y_{t-1}' \quad y_{t-2}' \quad \dots y_{t-p}']'$ is $(np+1) \times 1$, $\beta_t = vec([\,\Phi_{0,t} \quad \Phi_{1,t} \quad \dots \quad \Phi_{p,t}\,]')$;

$\sigma_t = [\,\sigma_{1,t} \quad \sigma_{2t}, \cdots \quad \sigma_{n,t}\,]'$ is a vector of reduced-form shock volatilities; $\Sigma_t^{\frac{1}{2}}$ ($\equiv diag(\sigma_t)$) is a diagonal matrix with σ_t on its main diagonal; and R is the correlation matrix.

In addition to the positive definite constraint, the correlation matrix R has diagonal elements fixed at one. This poses a potential difficulty in directly sampling the correlation matrix from an appropriate full conditional distribution. In order to circumvent this difficulty, we decompose $\Sigma_t^{\frac{1}{2}}$ into:

$$\Sigma_t^{\frac{1}{2}} = \tilde{\Sigma}_t^{\frac{1}{2}} \Sigma_0^{\frac{1}{2}}, \tag{3}$$

where $\Sigma_0^{\frac{1}{2}} = diag(\sigma_0)$, with $\sigma_0 = [\,\sigma_{1,0} \quad \sigma_{2,0} \quad \cdots \quad \sigma_{n,0}\,]'$ referring to a vector of initial volatilities; $\tilde{\Sigma}_t^{\frac{1}{2}} = diag(\tilde{\sigma}_t)$, with $\tilde{\sigma}_t = [\,\tilde{\sigma}_{1,t} \quad \tilde{\sigma}_{2,t} \quad \cdots \quad \tilde{\sigma}_{n,t}\,]'$ referring to a vector of the volatilities relative to their initial values ($\tilde{\sigma}_{i,t} = \frac{\sigma_{i,t}}{\sigma_{i,0}}$, $i = 1, 2, ..., n$). When the third equation in (2) is replaced by stationary log stochastic volatilities, σ_0 becomes a vector of steady-state or unconditional volatilities, and $\tilde{\sigma}_t$ becomes a vector of volatilities relative to their steady-state level.

Based on equation (3), we can reformulate our benchmark model in equation (2) in the following way:

Benchmark Model with Constant Correlation Matrix

$$\begin{aligned}
&y_t = (I_n \otimes x_t')\beta_t + u_t, \quad u_t \sim N(0, \Sigma_t^{\frac{1}{2}} R \Sigma_t^{\frac{1}{2}}) \equiv N(0, \tilde{\Sigma}_t^{\frac{1}{2}} G \tilde{\Sigma}_t^{\frac{1}{2}}),\\
&\beta_t = \beta_{t-1} + v_t, \quad v_t \sim i.i.d.N(0, \Omega_v),\\
&ln(\tilde{\sigma}_t) = ln(\tilde{\sigma}_{t-1}) + \eta_t, \quad ln(\tilde{\sigma}_t) = 0, \ \eta_t \sim i.i.d.N(0, \Omega_\eta),
\end{aligned} \tag{2'}$$

where $G = \Sigma_0^{\frac{1}{2}} R \Sigma_0^{\frac{1}{2}}$. This reformulation, especially for the error distribution, allows us to sample the G matrix from an appropriate inverse Wishart distribution, enabling us to avoid the challenging task of directly sampling the R matrix. Once the G matrix is sampled, we can obtain Σ_0 and R in the following way:

$$\Sigma_0 = diag(vecd(G)) \ \ and \ \ R = \Sigma_0^{-\frac{1}{2}} G \Sigma_0^{-\frac{1}{2}}. \tag{4}$$

where $diag(vecd(G))$ is a diagonal matrix with the diagonal elements of G in on the main diagonal.

In our extended model with a time-varying correlation matrix, we assume that the correlation matrix is subject to regime switching rather than continuous variation over time. We further assume that stationary log stochastic volatilities fluctuate around their regime-switching unconditional expectations. With regime shifts occurring concurrently, we have a regime-switching unconditional covariance matrix for the reduced-form errors. Then, we can extend our benchmark model in equation (2') in the following way: [7]

Extended Model with Regime-Switching Unconditional Covariance Matrix

$$\begin{aligned} y_t &= (I_n \otimes x_t')\beta_t + u_t, \quad u_t \sim N(0, \Sigma_t^{\frac{1}{2}} R_{S_t} \Sigma_t^{\frac{1}{2}}) \equiv N(0, \tilde{\Sigma}_t^{\frac{1}{2}} G_{S_t} \tilde{\Sigma}_t^{\frac{1}{2}}), \\ \beta_t &= \beta_{t-1} + v_t, \quad v_t \sim i.i.d.N(0, \Omega_v), \\ ln(\tilde{\sigma}_t) &= \Psi \; ln(\tilde{\sigma}_{t-1}) + \eta_t, \quad \eta_t \sim i.i.d.N(0, \Omega_\eta), \end{aligned} \tag{5}$$

where $G_{S_t} = \Sigma_{0,S_t}^{\frac{1}{2}} R_{S_t} \Sigma_{0,S_t}^{\frac{1}{2}}$ represents the regime-switching unconditional covariance matrix of the reduced-form error u_t; $\Psi = diag(\Psi_1, ..., \Psi_n)$, with $|\Psi_i| < 1$, $i = 1, 2, ..., n$; and S_t is a latent regime-indicator variable. We can specify the dynamics of S_t with the following transition probabilities:

$$Pr[S_t = j | S_{t-1} = i] = p_{ij}, \quad \sum_{j=1}^{M} p_{ij} = 1, \quad i = 1, 2, ..., M. \tag{6}$$

3. Outline of the MCMC Procedure for the Benchmark Model [8]

Throughout this section, we use the following notation to represent a vector of random variables or data, χ_κ, $\kappa = 1, 2, ..., t$:

$$\chi_{\leq t} = [\, \chi_1 \quad \chi_2 \quad \cdots \quad \chi_t \,]'.$$

[7] For a model with random-walk log stochastic volatilities and a regime-switching correlation matrix, the diagonal elements of the $\Sigma_0^{1/2}$ matrix in equation (3) represent the initial volatilities. The G_{S_t} matrix in equation (5) should be replaced by $G_{S_t} = \Sigma_0^{1/2} R_{S_t} \Sigma_0^{1/2}$, and we cannot avoid the challenging task of directly sampling the regime-switching correlation matrices.

[8] Extending the MCMC algorithm to the case of a regime-switching unconditional covariance matrix and stationary log volatilities in equations (5) and (6) is straightforward.

The key to successfully estimating the model specified in equations (2') lies in developing a procedure for sampling the stochastic volatilities of the reduced-form errors, conditional on data and other variates. A difficulty arises because the reduced-form errors are correlated with one another, which leads to correlations among their log squares, conditional on the stochastic volatilities.

In this section, we present an MCMC procedure with a focus on sampling the ordering-invariant multivariate log stochastic volatilities from the reduced-form errors. This is based on a multivariate extension of Kim, Shephard, and Chib's (1998) univariate model of stochastic volatility, which employs the mixture of normals approximation to the log-$\chi^2(1)$ distribution.

3.1. Sampling Log-Stochastic Volatilities ($\tilde{h}_{\leq T}$) and Ω_η Conditional on $ln(u_{\leq T})$ and R_ω

Keeping equation (3) in mind, we start with multiplying the first equation in (2') by $\Sigma_0^{-\frac{1}{2}}$ in order to obtain

$$\tilde{u}_t = \tilde{\Sigma}_t^{\frac{1}{2}} u_t^*, \quad u_t^* \sim i.i.d.N(0, R), \tag{7}$$

$$\Big(\ \tilde{u}_{i,t} = \tilde{\sigma}_{i,t} u_{i,t}^*, \quad u_{i,t}^* \sim i.i.d.N(0,1), \ i = 1, 2, ..., n, \ \Big)$$

where $\tilde{u}_t = \Sigma_0^{-\frac{1}{2}}(y_t - (I_n \otimes x_t')\beta_t)$, R is the correlation matrix, and $u_{i,t}^*$, $i = 1, 2, ..., n$, are correlated. We take the squares and then the natural logarithms of both sides of each equation in (7). Rearranging the resulting individual equations, along with the third equation in (2'), yields the following multivariate unobserved components model for the log stochastic volatilities:

$$\begin{aligned} ln(\tilde{u}_t^2) &= E(ln(u_t^{*2})) + 2\tilde{h}_t + \omega_t, \quad \omega_t \ \sim (0, \Omega_\omega), \\ \tilde{h}_t &= \tilde{h}_{t-1} + \eta_t, \quad \tilde{h}_0 = 0, \quad \eta_t \sim i.i.d.N(0, \Omega_\eta), \end{aligned} \tag{8}$$

where $\tilde{h}_t = ln(\tilde{\sigma}_t)$; $E(ln(u_t^{*2})) = -1.2704\iota$, with ι referring to an $n \times 1$ vector of ones; and $\omega_t = ln(u_t^{*2}) - E(ln(u_t^{*2}))$. Here, the elements of ω_t are correlated due to the correlation among the elements of u_t^*.

Since each element of ω_t follows a $log-\chi^2(1)$ distribution, the diagonal elements of Ω_ω are $\frac{\pi^2}{2}$. This allows us to decompose Ω_ω in equation (8) in the following way:

$$\Omega_\omega = \Omega_\omega^{*\frac{1}{2}} R_\omega \Omega_\omega^{*\frac{1}{2}}, \tag{9}$$

where $\Omega_\omega^{*\frac{1}{2}} = diag(\frac{\pi}{\sqrt{2}}, \frac{\pi}{\sqrt{2}}, \ldots, \frac{\pi}{\sqrt{2}})$ is an $n \times n$ diagonal matrix known a priori and R_ω is the correlation matrix of ω_t, which is a function of R. In what follows, we describe the details of our MCMC algorithm based on equations (8) and (9).

Following Kim, Shephard, and Chib (1998), we approximate the distribution for each row of ω_t in equation (8) using a mixture of normals, while explicitly accounting for the correlations among the rows of ω_t. Along with equation (9), this allows us to rewrite the first equation in (8) in the following way:

$$\begin{aligned}
&ln(\tilde{u}_t^2) = E(ln(u_t^{*2})) + 2\tilde{h}_t + \omega_t, \quad \omega_t \mid D_t \sim i.i.d.N(\mu^*_{\omega,D_t}, \Omega^{*\frac{1}{2}}_{\omega,D_t} R_\omega \Omega^{*\frac{1}{2}}_{\omega,D_t}),\\
&\mu^*_{\omega,D_t} = [\, \mu^*_{\omega,D_{1,t}} \quad \mu^*_{\omega,D_{2,t}} \cdots \quad \mu^*_{\omega,D_{n,t}} \,]',\\
&\Omega^*_{\omega,D_t} = diag(\sigma^{*2}_{\omega,D_{1,t}}, \sigma^{*2}_{\omega,D_{2,t}}, \cdots, \sigma^{*2}_{\omega,D_{n,t}}),\\
&Pr[D_{i,t} = \tau] = p^*_\tau, \quad \tau = 1, 2, ..., 7; \; i = 1, 2, ..., n,
\end{aligned} \tag{10}$$

where $D_t = [\, D_{1,t} \quad D_{2,t} \quad \ldots \quad D_{n,t} \,]'$ is an $n \times 1$ vector of mixture indicator variables; R_ω is the correlation matrix of ω_t; and the constants $\{\mu^*_{\omega,\tau}, \sigma^{*2}_{\omega,\tau}, p^*_\tau\}$ are known a priori. We assume that the rows of D_t are serially independent.

Due to the normality of ω_t conditional on D_t, equation (10) and the second equation in (8) allow us to sample $\tilde{h}_t$, $t = 1, 2, ..., T$, by using Carter and Kohn's (1994) procedure. The full conditional posterior distribution for $\tilde{h}_t$, $t = 1, 2, ..., T$, is given below:

$$\begin{aligned}
p(\tilde{h}_1, \tilde{h}_2, \ldots, \tilde{h}_T | \Omega_\eta, R_\omega, D_{\leq T}, ln(\tilde{u}^2_{\leq T})) \propto\; & p(\tilde{h}_T | \Omega_\eta, R_\omega, \mu^*_{\omega,D_T}, \Omega^*_{\omega,D_T}, ln(\tilde{u}^2_T))\\
& \times \prod_{t=T-1}^{1} p(\tilde{h}_t | \tilde{h}_{t+1}, \Omega_\eta, R_\omega, \mu^*_{\omega,D_t}, \Omega^*_{\omega,D_t}, ln(\tilde{u}^2_t)).
\end{aligned} \tag{11}$$

Once $\tilde{h}_t$, $t = 1, 2, ..., T$, are sampled, we can sample Ω_η based on the second equation in (8). The prior distribution we employ is: $\Omega_\eta \sim \; IW(W_\eta, d_\eta)$.

3.2. Sampling R_ω Conditional on R

For sampling R_ω conditional on R, we take advantage of the one-to-one correspondence between the moments of u_t^* and ω_t. For example, as derived by Harvey, Ruiz, and Shephard (1994), the following relation holds for the (i,j)-th elements of R and R_ω:

$$(R_\omega)_{ij} = \frac{2}{\pi^2} \sum_{g=1}^{\infty} \frac{(g-1)!}{(1/2)_g g} (R)_{ij}^{2g}, \quad i \neq j;\ i,j = 1,2,...,n, \tag{12}$$

where $(a)_g = a(a+1)\cdots(a+g-1)$.

We approximate equation (12) using simulations. We first generate K sets of $n \times 1$ random vectors z_i, for $i = 1,2,...,K$, from $z_i \sim i.i.d.N(0,R)$. We then calculate $\tilde{z}_i = ln(z_i^2)$, for $i = 1,2,...,K$, and use these to estimate and approximate the correlation matrix R_ω implied by R. We set K to be sufficiently large (e.g., $K = 100,000$).

3.3. Sampling Mixture Indicator Variables $D_{\leq T}$ ($\equiv [\, D_{1,\leq T} \quad D_{2,\leq T} \quad \dots \quad D_{n,\leq T}\,]$)

Conditional on D_t, the first equation in (10) can be rewritten as

$$ln(\tilde{u}_t^2) = -1.2704\iota + 2\tilde{h}_t + \mu_{\omega,D_t}^* + \Omega_{\omega,D_t}^{*\frac{1}{2}} \omega_t^*,\ \omega_t^* \ \sim i.i.d.N(0,R_\omega). \tag{13}$$

Here, as the rows of ω_t^* are correlated, the rows of D_t are correlated conditional on $ln(\tilde{u}_t^2)$ and $\tilde{h}_t$, even though they are unconditionally uncorrelated. We thus have to sample the rows of D_t jointly or sample each row from the full conditional distribution using the Gibbs sampling approach.

Based on equation (13), we can derive the following full conditional posterior distribution of $D_{i,\leq T}$, $i = 1,2,...,n$:

$$\begin{aligned} & p(D_{i,t} \mid ln(\tilde{u}_t^2), \tilde{h}_t, R_\omega, D_{\neq i,t}) \\ & \quad \propto p(D_{i,t}, ln(\tilde{u}_t^2) \mid \tilde{h}_t, R_\omega, D_{\neq i,t}) \\ & \quad = p(D_{i,t}) \times p(ln(\tilde{u}_t^2) \mid \tilde{h}_t, R_\omega, D_{\neq i,t}, D_{i,t}), \quad t = 1,2,...,T, \end{aligned} \tag{14}$$

where $D_{\neq i,t} = [\, D_{1,t} \quad \cdots D_{i-1,t} \quad D_{i+1,t} \quad \cdots \quad D_{n,t}\,]'$.

We initialize $D_{i,\leq T}$ for $i = 1,2,\dots,n$ based on their unconditional mixture probabilities from the last equation in (10). Then, at a particular iteration of the MCMC procedure, we can sample $D_{i,\leq T}$ sequentially for $i = 1,2,\dots,n$ based on the following probabilities:

$$Pr[D_{i,t} = \tau \mid ln(\tilde{u}_t^2), \tilde{h}_t, R_\omega, D_{\neq i,t}]$$
$$= \frac{p_\tau^* \times p(ln(\tilde{u}_t^2) \mid \tilde{h}_t, R_\omega, D_{\neq i,t}, D_{i,t} = \tau)}{\sum_{\tau=1}^{7} p_\tau^* \times p(ln(\tilde{u}_t^2) \mid \tilde{h}_t, R_\omega, D_{\neq i,t}, D_{i,t} = \tau)}, \; t = 1, 2, ..., T, \tag{15}$$

where p_τ^* represents the unconditional mixture probability in the last equation of (10).

3.4. The Rest of the MCMC Procedure

Sampling $\beta_{\leq T}$ and Ω_v Conditional on $\tilde{\Sigma}_{\leq T}$ and G

This step is based on the first two equations in (2'). Conditional on $\tilde{\Sigma}_{\leq T}$, Ω_v, G, and data, it is straightforward to sample $\beta_{\leq T}$ by employing the Carter and Kohn (1994) procedure. Once $\beta_{\leq T}$ is sampled, we can sample Ω_v conditional on $\beta_{\leq T}$. The prior distribution we employ is: $\Omega_v \sim IW(W_v, d_v)$.

Sampling G (and thus, R and Σ_0)

This step is based on the following distribution, which is obtained by multiplying the first equation in (2') by $\tilde{\Sigma}_t^{-\frac{1}{2}}$:

$$e_t \sim i.i.d.N(0, G), \tag{16}$$

where $e_t = \tilde{\Sigma}_t^{-\frac{1}{2}} u_t = \tilde{\Sigma}_t^{-\frac{1}{2}}(y_t - (I_n \otimes x_t')\beta_t)$. The prior distribution we employ is: $G \sim IW(W_G, d_G)$. Once G is sampled, we can obtain Σ_0 and R from equation (4).

4. Application #1: Forecasting Performance of the Benchmark Model

Arias, Rubio-Ramírez, and Shin (2023) (hereafter, ARS2023) demonstrate that the out-of-sample forecasting performance of Primiceri's (2005) model depends on the ordering of the variables. In particular, they show that when the object of interest is density forecasting, the differences in forecasting performance across orderings are noticeable and persistent.

In their forecasting exercise, they estimate the model in equation (1) with two lags ($p = 2$), using four quarterly variables: real GDP growth, inflation (based on the Core PCE Price Index), the 3-month T-bill rate, and the unemployment rate for the period from 1970Q3 to 2016Q4. They use 40 pre-sample observations covering the period from 1960Q3

to 1970Q2 to calibrate the priors. For each of the 24 possible orderings of the variables, they recursively estimate the model and generate one-, four-, and eight-quarter-ahead forecasts ($h = 1,\ 4,\ 8$) over 120 quarters ($c = 1, 2, \ldots, 120$) starting in 1987Q1. Their latest forecast is for 2018Q4 when $c = 120$ and $h = 8$.

In this section, we use the same data set and procedure as in ARS2023 to evaluate the out-of-sample forecasting performance of our benchmark ordering-invariant model in equation (2'), which features a constant correlation matrix and random-walk log stochastic volatilities. We compare the forecasting performance of our benchmark model against Primiceri's (2005) ordering-dependent model in equation (1) by using log predictive scores (LPSs) for density forecasts and RMSEs for point forecasts.

The priors and hyperparameters of the parameters common to both Primiceri's (2005) model and our benchmark model are set to be the same as in ARS2023. Specifically, the priors and hyperparameters for the initial coefficients vector (β_0) and the covariance matrix of the shocks to the coefficients (Ω_v) are specified as follows:

$$\begin{aligned}
&\beta_0 \sim N(\hat{\beta}_{pre}, Cov(\hat{\beta}_{pre}) \times 4), \\
&\Omega_v \sim IW(k_v^2 \times Cov(\hat{\beta}_{pre}) \times 40,\ 40),\ with\ k_v = 0.01,
\end{aligned}$$

where $\hat{\beta}_{pre}$ is the OLS estimator of the constant β coefficient for the pre-sample and $Cov(\hat{\beta}_{pre})$ is the covariance matrix of $\hat{\beta}_{pre}$. The priors and hyperparameters for the reduced-form error distribution specific to our model are specified as follows: [9]

$$\begin{aligned}
&\Omega_\eta \sim IW(k_\eta^2 \times diag(vecd(\hat{\Sigma}_{pre}))^{\frac{1}{2}} \times 5,\ 5)\ with\ k_\eta = 0.1, \\
&G \sim IW(k_G^2 \times diag(vecd(\hat{\Sigma}_{pre})) \times 5,\ 5)\ with\ k_G = 1,
\end{aligned}$$

where $\hat{\Sigma}_{pre}$ is the covariance matrix of the reduced-form errors estimated from the pre-sample, and $diag(vecd(\hat{\Sigma}_{pre}))$ is a diagonal matrix with the diagonal elements of $\hat{\Sigma}_{pre}$ on its main diagonal. We note that, unlike the tight prior for the covariance matrix of the shocks to the coefficients (Ω_v), the priors for Ω_η and G are not very informative.

As in ARS2023, we generate 110,000 MCMC parameter draws, using the first 10,000 of these as a burn-in period and storing every 10th draw from the chain. Our forecasts are

[9] Unlike for Primiceri's (2005) model, $\hat{\Sigma}_{pre}$ obtained from the pre-sample is informative about Ω_η and G for our model.

based on the resulting 10,000 MCMC parameter draws.

The results for density forecasts are presented in Table 1. For joint density forecasts, our model's LPS is almost the same as or higher than the highest LPS from Primiceri's (2005) model for $h = 1$ and $h = 4$, while it is somewhat lower for $h = 8$. For individual density forecasts, in eight out of twelve cases, our model's LPSs are higher than or slightly lower than the highest values from Primiceri's (2005) model. In the other four cases, our model's LPSs are close to the median values from Primiceri's (2005) model. To summarize, the density forecasting performance of our benchmark model is mostly comparable to the best performance of Primiceri's (2005) model across orderings.

The results for point forecasts are presented in Table 2. In most cases, the RMSEs from our model are the same as or lower than the median RMSEs from Primiceri's (2005) model. In only three out of twelve cases are the RMSEs from our model slightly higher than the median RMSEs from Primiceri's (2005) model. To summarize, the point forecasting performance of our benchmark model is comparable to the median performance of Primiceri's (2005) model across orderings.

5. Application #2: An Investigation of the Changing Business Cycle Dynamics in the U.S. Economy and Its Implications for the Hysteresis Effect

The goal of this section is to investigate how the relative importance of permanent and transitory shocks has changed over time in the U.S. economy. For this, we employ Blanchard and Quah's (1989) bivariate VAR model of real GDP growth and the unemployment rate, extended to incorporate time-varying coefficients and stochastic volatilities.

5.1. Empirical Model Specification

Let us consider the following situation in which the aggregate demand and aggregate supply shocks are correlated:

$$\begin{bmatrix} \tilde{\varepsilon}_t^{AS} \\ \tilde{\varepsilon}_t^{AD} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ d_t & 1 \end{bmatrix} \begin{bmatrix} \hat{\varepsilon}_t^{P} \\ \hat{\varepsilon}_t^{T} \end{bmatrix}, \quad \begin{bmatrix} \hat{\varepsilon}_t^{P} \\ \hat{\varepsilon}_t^{T} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{1,t}^{*2} & 0 \\ 0 & \sigma_{2,t}^{*2} \end{bmatrix} \right), \tag{17}$$

$$(\tilde{\varepsilon}_t = D_t \hat{\varepsilon}_t, \ \hat{\varepsilon}_t \sim N(0, \Sigma_t^*),)$$

where $\tilde{\varepsilon}_t^{AS}$ and $\tilde{\varepsilon}_t^{AD}$ refer to aggregate supply and aggregate demand shocks, respectively and $\hat{\varepsilon}_t^{P}$ and $\hat{\varepsilon}_t^{T}$ refer to orthogonalized permanent and transitory shocks, respectively. Equation (17) suggests that aggregate demand shock is a function of both permanent and transitory shocks, while the aggregate supply shock is a function of only the permanent shock.

Here, the $\sigma_{1,t}^*$ term represents the the volatility of the permanent shock. The d_t parameter represents the extent to which the aggregate demand shock has permanent effects, i.e., it measures the hysteresis effect. The $\sigma_{2,t}^*$ term represents the volatility of the transitory shock.

In this situation, we consider the following encompassing model:

$$\begin{aligned} &y_t = \Phi_{0,t} + \Phi_{1,t} y_{t-1} + \Phi_{2,t} y_{t-2} + \cdots + \Phi_{p,t} y_{t-p} + u_t, \quad u_t \sim N(0, \tilde{\Sigma}_t^{\frac{1}{2}} G_{S_t} \tilde{\Sigma}_t^{\frac{1}{2}}), \\ &\tilde{h}_t = \Psi \tilde{h}_{t-1} + \eta_t, \quad \eta_t \sim i.i.d.N(0, \Omega_\eta), \end{aligned} \tag{18}$$

where y_t is a 2×1 vector that consists of real GDP growth and the unemployment rate. The coefficients follow random walk processes. As before, $\tilde{h}_t = \ln(\tilde{\sigma}_t)$; $\tilde{\sigma}_t$ is a vector that consists of the diagonal elements of $\tilde{\Sigma}_t^{\frac{1}{2}}$; and $\Psi = diag(\Psi_1, \Psi_2)$. In the case of a constant correlation matrix and random-walk log stochastic volatilities, we have $G_{S_t} = G$ and $\Psi = I_2$. For a regime-switching covariance matrix for the reduced-form error, we specify the dynamics of the regime-indicator variable (S_t) as in equation (6).

The relationship between the reduced-form error (u_t) and the normalized structural error (ε_t) is given below:

$$u_t = B_t^{-1} \varepsilon_t, \quad \varepsilon_t \ \sim \ i.i.d.N(0, I_n), \tag{19}$$

where $\varepsilon_t = \Sigma_t^{*-\frac{1}{2}} \hat{\varepsilon}_t$, and

$$B_t^{-1} = A_t^{-1} D_t \Sigma_t^{*\frac{1}{2}} = \frac{1}{1 - a_{12,t} a_{21,t}} \begin{bmatrix} (1 + a_{12,t} d_t)\sigma_{1,t}^* & a_{12,t}\sigma_{2,t}^* \\ (a_{21,t} + d_t)\sigma_{1,t}^* & \sigma_{2,t}^* \end{bmatrix}, \tag{20}$$

where the negative of $a_{ij,t}$ is the (i,j)-th off-diagonal element of A_t. [10] Unlike in equation (1), the A_t matrix in equation (20) is unconstrained except for the units on the main diagonal.

[10] The B_t matrix can be derived as

$$B_t = \begin{bmatrix} 1/\sigma_{1,t}^* & -a_{12,t}/\sigma_{1,t}^* \\ -(d_t + a_{21,t})/\sigma_{2,t}^* & (1 + a_{12,t} d_t)/\sigma_{2,t}^* \end{bmatrix}. \tag{20'}$$

Once B_t^{-1}, and thus B_t, is identified, the $a_{12,t}$ parameter can be obtained by dividing the (1,2) element of B_t^{-1} by its (2,2) element. The inverse of the (1,1) element of B_t in equation (20') in Footnote 10 provides the volatility of the permanent shock through the aggregate supply shock ($\sigma^*_{1,t}$). Even though the volatility of the transitory shock ($\sigma^*_{2,t}$) is not separately identified, we can obtain the following ratio, by dividing the (1,1) element of B_t^{-1} by its (2,2) element: [11]

$$r_t = \frac{(1 + a_{12,t}d_t)\sigma^*_{1,t}}{\sigma^*_{2,t}}, \tag{21}$$

where r_t represents the ratio of the total permanent shock volatility to the transitory shock volatility. In the numerator of equation (21), the term $a_{12,t}d_t \times \sigma^*_{1,t}$ can be interpreted as the portion of the total permanent shock volatility attributable to the hysteresis effect.

5.2. Identification of the B_t^{-1} Matrix

For identification of the B_t^{-1} matrix, we assume that the transitory shock has no long-run impact on the level of real output. This is equivalent to assuming that the long-run response matrix $\theta_t(1)$, shown below, is lower triangular:

$$\theta_t(1) = \sum_{j=1}^{\infty} \frac{\partial y_{t+j}}{\partial \varepsilon_t}, \quad t = 1, 2, ..., T. \tag{22}$$

That is, we identify the B_t^{-1} matrix by applying the Blanchard and Quah (1989) procedure at each point in time.

The model in equation (18) can be written in the following companion form, ignoring the intercept terms:

$$\tilde{y}_t = F_t \tilde{y}_{t-1} + \tilde{u}_t, \tag{23}$$

where $\tilde{y}_t = [\, y_t' \quad y_{t-1}' \quad \cdots \quad y_{t-p+1}' \,]'$, $\tilde{u}_t = [\, \varepsilon_t' B_t^{-1'} \quad 0' \quad \cdots \quad 0' \,]'$, and F_t can be appropriately defined as a function of $\Phi_{1,t}, \ldots, \Phi_{p,t}$. We then have

$$\tilde{y}_{t+j} = \tilde{u}_{t+j} + F_{t+j}\tilde{u}_{t+j-1} + (\prod_{i=j-1}^{j} F_{t+i})\tilde{u}_{t+j-2} + \ldots + (\prod_{i=1}^{j} F_{t+i})\tilde{u}_t + (\prod_{i=0}^{j} F_{t+i})\tilde{y}_{t-1}, \tag{24}$$

[11] Equation (21) can also be obtained by dividing (2,2) element of B_t by its (1,1) element.

from which we obtain the following impulse-response function:

$$\begin{aligned}\frac{\partial y_{t+j}}{\partial \varepsilon_t} &= (\prod_{i=1}^{j} E_t[F_{t+i}])_{(1:2,1:2)} B_t^{-1} \\ &= (F_t^j)_{(1:2,1:2)} B_t^{-1},\end{aligned} \tag{25}$$

as $E_t[F_{t+i}] = F_t$ for all i and $(F_t^j)_{(1:2,1:2)}$, for example, is the first 2×2 block of F_t^j. We thus have the following results:

$$\begin{aligned}&\theta_t(1) = \sum_{j=0}^{\infty} \frac{\partial y_{t+j}}{\partial \varepsilon_t} = ((I_{2p} - F_t)^{-1})_{(1:2,1:2)} B_t^{-1}, \\ &\theta_t(1)\theta_t(1)' = ((I_{2p} - F_t)^{-1})_{(1:2,1:2)} \tilde{\Sigma}_t^{\frac{1}{2}} G_{S_t} \tilde{\Sigma}_t^{\frac{1}{2}} ((I_{2p} - F_t)^{-1})'_{(1:2,1:2)},\end{aligned} \tag{26}$$

as $B_t^{-1} B_t^{-1'} = \tilde{\Sigma}_t^{\frac{1}{2}} G_{S_t} \tilde{\Sigma}_t^{\frac{1}{2}}$.

Since $\theta_t(1)$ is lower triangular, it can be obtained by the Cholesky decomposition of the right-hand side of the second equation in (26). Given $\theta_t(1)$, the B_t^{-1} matrix can be obtained from the first equation in (26).

5.3. Empirical Results

5.3.1. Data and Priors

We use quarterly data on real GDP growth and the unemployment rate obtained from the Federal Reserve Economic Data Base (FRED), covering the sample period from 1960Q2 to 2019Q4. Monthly unemployment rate data are averaged to generate quarterly observations. The data in our main sample are plotted in Figure 1, with shaded areas representing NBER recession periods.

We estimate four alternative models: i) a model with constant correlation matrix and random walk log stochastic volatilities; ii) a model with constant correlation matrix and stationary log stochastic volatilities; iii) a model with two-state regime-switching covariance matrix with stationary log stochastic volatilities; and iv) a model with three-state regime-switching covariance matrix with stationary log stochastic volatilities.

To calibrate the priors, we use 40 pre-sample observations covering the period from 1950Q2 to 1960Q1. We employ the following priors and hyperparameters for the parameters

common to all four models under consideration:

$$\beta_0 \sim N(\hat{\beta}_{pre}, Cov(\hat{\beta}_{pre}) \times 4),$$
$$\Omega_v \sim IW(k_v^2 \times Cov(\hat{\beta}_{pre}) \times 40,\ 40),\ with\ (k_{v,a}, k_{v,b}) = (0.03, 0.04),$$
$$\Omega_\eta \sim IW(k_\eta^2 \times diag(vecd(\hat{\Sigma}_{pre}))^{\frac{1}{2}} \times 5,\ 5),\quad with\ k_\eta = 0.1,$$

where $\hat{\beta}_{pre}$ is the OLS estimator of the constant β coefficient for the pre-sample; $Cov(\hat{\beta}_{pre})$ is the covariance matrix of $\hat{\beta}_{pre}$; $k_{v,a}$ is the value of k_v for the VAR slope coefficients and $k_{v,b}$ is the value of k_v for the VAR intercept coefficients. We specify that the variances of the shocks to the intercept coefficients are higher than those to the slope coefficients.

For models with constant correlation matrix, we employ the following prior for the G matrix:

$$G \sim IW(k_G^2 \times \hat{W} \times 5,\ 5),\ with\ \hat{W} = diag(vecd(\hat{\Sigma}_{pre})),$$
$$k_G = 1\ for\ r.w.\ log\ volatilities,\ and\ 0.5\ for\ stationary\ log\ volatilities.$$

For models with stationary log stochastic volatilities, we additionally employ the priors for the persistence parameters of the volatility: $\Psi_i \sim N(0.9, 1)$, $i = 1, 2$.

For a model with two-state regime-switching covariance matrix with stationary log stochastic volatilities, we employ the following additional priors: $G_1 \sim IW(k_G^2 \times \hat{W} \times 5,\ 5)$; $G_2 \sim IW(k_G^2 \times (0.5 \times \hat{W}) \times 5,\ 5)$; $[\,p_{11} \quad p_{12}\,]' \sim Dir(19, 1)$; $[\,p_{21} \quad p_{22}\,]' \sim Dir(1, 19)$, where $Dir(\cdot)$ refers to the Dirichlet distribution. For a model with three-state regime-switching covariance matrix with stationary log stochastic volatilities, we employ the following additional priors: $G_3 \sim IW(k_G^2 \times (0.3 \times \hat{W}) \times 5,\ 5)$; $[\,p_{11} \quad p_{12} \quad p_{13}\,]' \sim Dir(19, 0.5, 0.5)$; $[\,p_{21} \quad p_{22} \quad p_{23}\,]' \sim Dir(0.5, 19, 0.5)$; $[\,p_{31} \quad p_{32} \quad p_{33}\,]' \sim Dir(0.5, 0.5, 19)$. Note that we employ informative priors for the transition probabilities so that regime switching captures the low frequency variations in the log stochastic volatilities.

5.3.2. Empirical Results: Evidence of Changing Business Cycle Dynamics in the U.S. Economy

For each of the models under consideration, we use four lags ($p = 4$). The MCMC simulations are based on 50,000 iterations of the Gibbs sampler, discarding the first 10,000 for convergence. As a basis for our model selection, we employ the WAIC (Watanabe-Akaike

Information Criterion). A two-state regime switching model with stationary log stochastic volatilities resulted in the lowest value for the WAIC. In particular, this model is strongly preferred over the model with constant correlation matrix and random-walk log stochastic volatilities. [12] Empirical results presented in this section are based on the two-state regime-switching model.

Permanent Shock Volatility and the Ratio r_t

Putting aside the Great Recession period in the left panel of Figure 2, there appears to be a structural break in the volatility of the permanent shock through the aggregate supply shock ($\sigma^*_{1,t}$) , moving toward a lower level since the mid-1980s. There seems to be no distinct increasing or decreasing pattern after the structural break.

The right panel of Figure 2 shows that the ratio of total permanent shock to transitory shock volatilities (r_t) defined in equation (21) exhibits no distinct increasing or decreasing pattern for the entire sample period. It seems that this ratio fluctuates around a steady-state mean. In the absence of the hysteresis effect ($d_t = 0$), this would mean that the ratio of the aggregate supply to aggregate demand shock volatilities fluctuate around its steady-state mean.

Impulse-Responses

We observe significant changes in the impulse response patterns over time. The upper panels of Figure 3 show that while the short-run (0 quarters ahead) and medium-run (4 quarters ahead) responses of real output to transitory and permanent shocks were of similar magnitude prior to the mid-1980s, they have diverged since then. In particular, the responses of output to transitory shocks in both the short and medium runs have decreased considerably since the mid-1980s.

The lower left panel of Figure 3 shows that a positive permanent shock prior to the mid-1990s generally resulted in a short-run increase in the unemployment rate, followed by a decrease after a few quarters. In the long run, the permanent shock has no effect on the unemployment rate. [13] However, since the mid-1990s, particularly during the Great

[12] We observe that, once regime switching is allowed in the mean of the log stochastic volatilities, their persistence is significantly reduced compared to random-walk log stochastic volatilities.

[13] This pattern resembles the response of the unemployment rate to a positive aggregate

Recession period, the responses of the unemployment rate to permanent shocks resemble those to transitory shocks, as presented in the lower right panel of Figure 3. Specifically, the short-run response of the unemployment rate to permanent shocks has turned negative.

Variance Decompositions

The variance decompositions of the h-quarter-ahead prediction errors ($h = 1, 4, 40$) are depicted in Figure 4. The variances of the prediction errors for real output and the unemployment rate due to permanent shocks (through both aggregate supply shock and the hysteresis effect) have been increasing since the mid-1980s, although the evidence is less compelling for the unemployment rate.

5.3.3. Interpretation of Empirical Results and Implications for the Hysteresis Effect

In order to investigate the source of our empirical results, which indicate an increasing relative importance of permanent shocks in the dynamics of real output since the mid-1980s, we consider the following equation for the variance of the 1-step-ahead prediction error for real output due to the permanent shock:

$$\begin{aligned} V_t &= \frac{(1 + a_{12,t} d_t)^2 \sigma_{1,t}^{*2}}{(1 + a_{12,t} d_t)^2 \sigma_{1,t}^{*2} + a_{12,t}^2 \sigma_{2,t}^{*2}} \\ &= \frac{r_t^2}{r_t^2 + a_{12,t}^2}, \end{aligned} \tag{27}$$

where r_t, as defined in equation (21), is the ratio of total permanent shock volatility to transitory shock volatility.

Given that r_t exhibits a stable pattern, as depicted in the right panel of Figure 2, equation (27) suggests that the decreasing $a_{12,t}$ parameter shown in Figure 5 is an important source of the increasing V_t since the mid-1980s. Note that this also resulted in a decreasing response of real output to the transitory shock, as depicted in the upper right panel of Figure 3. In what follows, we present indirect evidence showing that the hysteresis effect ($d_t > 0$) has also played an important role in the increasing V_t.

First, focusing on the short-run fluctuations in the ratio r_t, the right panel of Figure 2

supply shock reported in Blanchard and Quah (1989), who estimated a time-invariant VAR model using a sample covering the period 1950Q2 to 1987Q4.

shows that all recessions are associated with a similar decline in this ratio. For example, the decrease in this ratio during the Great Recession period is no larger than that for other aggregate-supply-oriented recessions, especially in the 1970s. Considering that the Great Recession was primarily caused by a negative aggregate demand shock with increased volatility, this cannot be justified in the absence of the hysteresis effect. Without the hysteresis effect, we can easily conjecture a significant decrease in the ratio of the aggregate supply shock to aggregate demand shock volatilities ($\frac{\sigma^*_{1,t}}{\sigma_{2,t}}$). [14] Then, equation (21) and the second line of equation (27) suggest that during the Great Recession period, we would have observed a much larger decrease in r_t, and thus in V_t, than we actually observe from our empirical results.

Second, additional indirect evidence of the hysteresis effect can be found in the Beveridge-Nelson (1981) trend-cycle decomposition of real GDP, as depicted in Figure 6.A. Focusing on the trend component in the right panel, we observe a change in its dynamics since the 1990 recession. Recessions since then have had permanent effects on the level of real output. In particular, a significant portion of the decline in real output is attributed to a decline in the trend component. This again demonstrates the hysteresis effect, considering that the cause of the Great Recession was mostly aggregate-demand oriented.

Third, the change in the response pattern of the unemployment rate to a permanent shock, as shown in the lower left panel of Figure 3, may also be due to the hysteresis effect. Since the mid-1990s, this pattern resembles the response to a transitory shock, with the short-run response to a permanent shock changing its sign from positive to negative.

In relation to the existing literature, Cerra, Fatás, and Saxena (2023) and the references therein provide empirical evidence supporting the hysteresis effect interpretation of our empirical results. Based on the cross-sectional analysis, they focus on the permanent effects of negative demand shocks, such as the Asian crisis or the Global Financial Crisis, on real output.

5.3.4. Sensitivity of the Results to Alternative Priors

[14] Without the hysteresis effect, the transitory shock is the same as the aggregate demand shock in our model.

We first investigated the sensitivity of the results to the prior distributions for Ω_η and G. Additional sets of corresponding hyperparameters k_η and k_G we considered, while keeping all the other hyperparameters unchanged, are: $(k_\eta, k_G) = (0.05, 1),\ (0.1, 0.5),\ (0.05, 0.5)$. The results were not very sensitive, which makes sense because the implied prior distributions for Ω_η and G are not very informative given the degrees of freedom of 5.

The results were somewhat sensitive to the hyperparameters of the prior distribution for Ω_v, even though the qualitative results remained unchanged. This sensitivity arises because the implied prior distribution for Ω_v is very informative given the degrees of freedom of 40. We considered additional sets of corresponding hyperparameters $k_{v,a}$ and $k_{v,b}$, while keeping all other hyperparameters unchanged: $(k_{v,a}, k_{v,b}) = (0.02, 0.03),\ (0.01, 0.02)$. As the values of these two parameters decrease, the variance decomposition pattern in Figure 2 became flatter, while still indicating an increasing relative importance of the permanent shock.

Lastly, when we set $k_{v,a} = k_{v,b}$, the qualitative results are the same as before. Figure 6.B depicts the trend-cycle decomposition of real output when we set $(k_{v,a}, k_{v,b}) = (0.03, 0.03)$. We have much higher uncertainty associated with the cyclical component during the Great Recession. Subject to this higher uncertainty, most of the decline in real output during this period is associated with a drop in the trend component, providing stronger evidence in favor of the hysteresis effect than in the case where $k_{v,a} < k_{v,b}$.

6. Concluding Remarks

In this paper, we present a reduced-form TVP-VAR model that is ordering-invariant and easy to estimate. All we need for the estimation of our model is a simple and straightforward Gibbs sampling approach. And yet, our benchmark model, featuring a constant correlation matrix and random-walk log stochastic volatilities, provides density forecasting performance mostly comparable to the best performance of Primiceri's (2005) model across orderings, even though the point forecasting performance matches its median performance.

Recently, Hafner and Wang (2023) propose a new model for the time-varying correlation matrix, where the dynamics are driven by the likelihood score with respect to the matrix logarithm of the correlation matrix. Arias, Rubio-Ramrez, and Shin (2023) also propose the

elliptical slice sampling procedure for sampling the random-walk matrix logarithm of the time-varying correlation matrix. Incorporating the continuously time-varying correlation matrix into our model along these lines would be a future research agenda.

References

[1] Archakov, I. and P.R. Hansen, 2021, "A New Parameterization of Correlation Matrices," *Econometrica,* 89 (4), 16991715.

[2] 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.

[3] Arias, J.E., J.F. Rubio-Ramírez, M. Shin, and D.F. Waggoner, 2024, "Inference Based on Time-Varying SVARs Identified with Sign Restrictions," Federal Reserve Bank of Philadelphia, Working Paper 24-05.

[4] Asai, M. and M. McAleer, 2009, "The structure of dynamic correlations in multivariate stochastic volatility models," *Journal of Econometrics, 150(2), 182-192.*

[5] Bertsche, D. and R. Braun (2022), "Identification of Structural Vector Autoregressions by Stochastic Volatility," *Journal of Business & Economic Statistics,* 40, 328341.

[6] Beveridge, S. and C.R. Nelson, 1981, "New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, 7, 151-174.

[7] Benati, L., 2008, 'The Great Moderation in the United Kingdom," *Journal of Money, Credit and Banking*, 40(1), 121-147.

[8] Blanchard, Olivier J. and Danny Quah, 1989, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *The American Economic Review*, Vol. 79, No. 4, 655-673.

[9] Bognanni, M., 2018, "A Class of Time-Varying Parameter Structural VARs for Inference Under Exact or Set Identification," Federal Reserve Bank of Cleveland Working Paper 1 (18-11), 161.

[10] Carriero, A., T. E. Clark, and M. Marcellino, 2019, "Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-conjugate Priors," *Journal of Econometrics*,

212 (1), 137-154.

[11] Carter, C. K. and R. Kohn, 1994, "On Gibbs Sampling for State Space Models," *Biometrika*, Vol. 81, No. 3, 541-553.

[12] Cerra, V., A. Fatás, and S.C. Saxena, 2023, "Hysteresis and Business Cycles," *Journal of Economic Literature,* Vol. 61, No.1, 181-225.

[13] Chan, J.C.C., 2023, "Large Hybrid Time-Varying Parameter VARs," *Journal of Business & Economic Statistics, 41:3, 890-905.*

[14] Chan, J.C.C and E. Eisenstat, 2018, "Bayesian Model Comparison for Time-Varying Parameter VARs with Stochastic Volatility," *Journal of Applied Econometrics*, 33, 509-532.

[15] Chan, J.C.C., Koop, G. and X. Yu, 2024, "Large Order-Invariant Bayesian VARs with Stochastic Volatility," *Journal of Business and Economic Statistics*, 42:2, 825-837.

[16] Clark, T. E., and F. Ravazzolo, 2015, "Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics*, 30, 551-575.

[17] Cogley, T. and T. J. Sargent, 2001, "Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual,* Vol. 16 (2001), 331-373.

[18] Cogley, T. and T. J. Sargent, 2005, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US," *Review of Economic Dynamics,* 8, 262-302.

[19] D'Agostino, A., Gambetti, L., and D. Giannone, 2013, "Macroeconomic Forecasting and Structural Change," *Journal of Applied Econometrics*, 28, 82-101.

[20] Fischer, M. M., N. Hauzenberger, F. Huber, and M. Pfarrhofer, 2023, "General Bayesian Time-Varying Parameter Vector Autoregressions for Modeling Government Bond Yields," *Journal of Applied Econometrics,* 38(1), 69-87.

[21] Gali, J., and L. Gambetti, 2009, "On the Sources of the Great Moderation," *American Economic Journal: Macroeconomics*, 1(1), 26-57.

[22] Gambetti, L., and A. Musso, 2017, "Loan supply shocks and the business cycle," *Journal of Applied Econometrics*, 32, 764-782.

[23] Golosnoy, V., B. Gribisch, and R. Liesenfeld, 2012, "The Conditional Autoregressive Wishart Model for Multivariate Stock Market Volatility," *Journal of Econometrics*, 167, 211-223.

[24] Hafner, C.M. and L. Wang, 2023, "A Dynamic Conditional Score Model for the Log Correlation Matrix," *Journal of Econometrics,* 237, Article 105176.

[25] Harvey, A., E. Ruiz, and N. Shephard, 1994, "Multivariate Stochastic Variance Models," *The Review of Economic Studies*, Vol. 61, No. 2, pp. 247-264.

[26] Kim, S., N. Shephard, and S. Chib, 1998, "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, Vol. 65, No. 3, 361-393.

[27] Koop, G., and D. Korobilis, 2013, "Large Time-Varying Parameter VARs," *Journal of Econometrics*, 177, 185-198.

[28] Mumtaz, H. and K. Petrova, 2023, "Changing Impact of Shocks: A Time-Varying Proxy SVAR Approach," *Journal of Money, Credit, and Banking*, Vol. 55, No. 2-3, 635-654.

[29] Nelson, C.R. and C.I. Plosser, 1982, "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics,* 10, 139-162.

[30] Philipov, A., and M. Glickman, 2006, "Multivariate stochastic volatility via Wishart processes," *Journal of Business and Economic Statistics*, 24, 313-328.

[31] Pietro, E., S. Eickmeier, and M. Marcellino, 2016, "Time Variation in Macro-Financial Linkages," *Journal of Applied Econometrics*, 31(7), 1215-1233.

[32] Primiceri, Giorgio E., 2005, "Time Varying Structural Vector Autoregressions and Monetary Policy," *The Review of Economic Studies,* 72, 821-852.

[33] Shin, M. and M. Zhong, 2020, "A new approach to identifying the real effects of uncertainty shocks" *Journal of Business and Economic Statistics*, 38(2), 367-379.

[34] Wu, P. and G. Koop, 2023, "Estimating the Ordering of Variables in a VAR Using a Plackett-Luce Prior," *Economics Letters,* Vol. 230, September, 111247.

[35] Zheng, T., S. Ye, S., and Y. Hong, 2023, "Fast Estimation of a Large TVP-VAR Model with Score-Driven Volatilities," *Journal of Economic Dynamics and Control*, December 2023, Vol.157, Article 104762.

Table 1. Comparison of Out-of-Sample Forecasting Performances: Primiceri's (2005) model vs. Our Ordering-Invariant Model [Log Predictive Scores for density forecasts]

1-quarter-ahead

	Primiceri's (2005) Model (from Arias et al. (2023), Table 2) [minimum, median, maximum]	Our Model
Joint	[-379.44, -359.70, -350.06]	-350.09
Real Output	[-281.65, -278.53, -274.66]	-276.98
Inflation	[-113.86, -111.91, -111.31]	-110.19
3-Month T-Bill	[-28.76, -14.85, -10.62]	-16.13
Unemployment	[21.01, 27.18, 30.15]	30.86

4-quarter-ahead

	Primiceri's (2005) Model (from Arias et al. (2023), Table 2) [minimum, median, maximum]	Our Model
Joint	[-773.92, -750.89, -733.46]	-720.07
Real Output	[-292.67, -290.13, -286.59]	-287.18
Inflation	[-152.88, -148.31, -146.31]	-145.07
3-Month T-Bill	[-193.21, -184.32, -180.87]	-184.19
Unemployment	[-141.95, -134.47, -126.13]	-127.85

8-quarter-ahead

	Primiceri's (2005) Model (from Arias et al. (2023), Table 2) [minimum, median, maximum]	Our Model
Joint	[-969.13, -940.03, -917.53]	-927.91
Real Output	[-294.80, -290.01, -287.51]	-287.09
Inflation	[-182.95, -177.93, -174.56]	-175.57
3-Month T-Bill	[-264.91, -258.94, -256.47]	-260.79
Unemployment	[-220.76, -214.12, -206.55]	-213.73

Table 2. Comparison of Out-of-Sample Forecasting Performances: Primiceri's (2005) Model vs. Our Ordering-Invariant Model [RMSEs for point forecasts]

	Primiceri's (2005) Model (from Arias et al. (2023), Table 1) [minimum, median, maximum]	Our Model
<u>1-quarter-ahead</u>		
Real Output	[2.49, 2.52, 2.58]	2.55
Inflation	[0.59, 0.60, 0.60]	0.61
3-Month T-Bill	[0.32, 0.33, 0.34]	0.33
Unemployment	[0.19, 0.20, 0.20]	0.20
<u>4-quarter-ahead</u>		
Real Output	[2.60, 2.66, 2.72]	2.63
Inflation	[0.76, 0.78, 0.80]	0.79
3-Month T-Bill	[1.13, 1.15, 1.18]	1.14
Unemployment	[0.79, 0.81, 0.82]	0.81
<u>8-quarter-ahead</u>		
Real Output	[2.44, 2.53, 2.60]	2.51
Inflation	[0.85, 0.89, 0.95]	0.91
3-Month T-Bill	[1.86, 1.91, 1.94]	1.93
Unemployment	[1.40, 1.44, 1.47]	1.41

Figure 1. Data: Real GDP Growth and Unemployment Rate

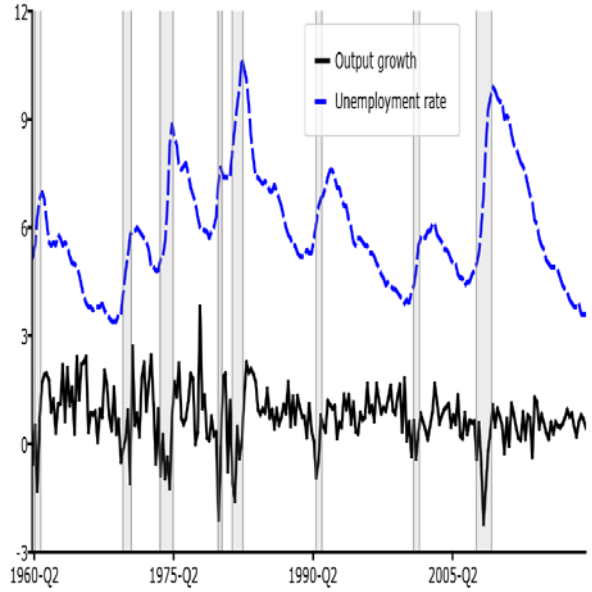
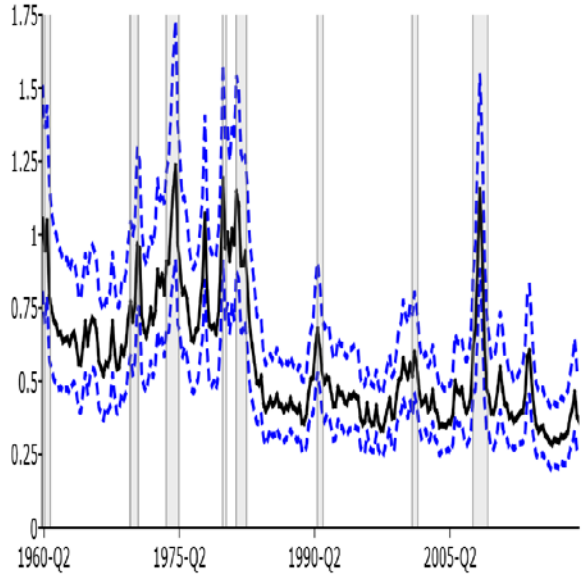
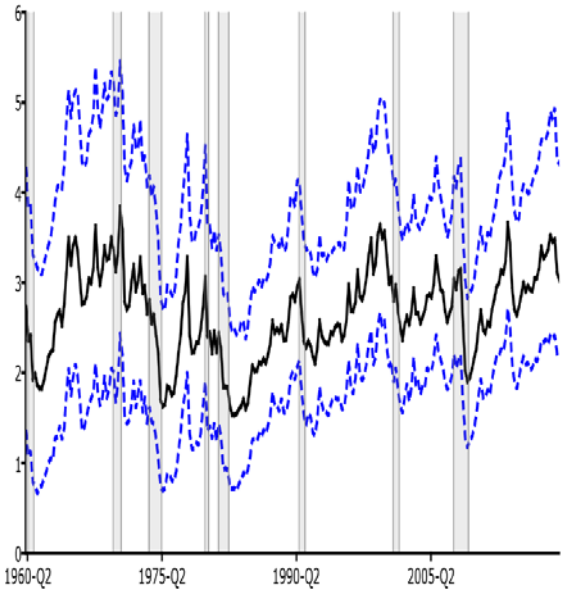


Figure 2. Permanent Shock Volatility ($\sigma^*_{1,t}$), and the Ratio of Total Permanent Shock Volatility to Transitory Shock Volatility ($r_t = (1 + a_{12,t}d_t)\,\sigma^*_{1,t} \;/\sigma^*_{2,t}$) [Posterior Medians with 64% Posterior Bands]

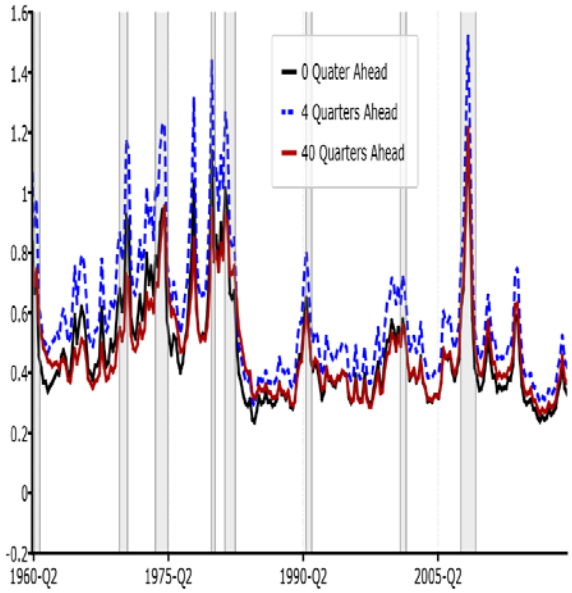


Permanent Shock Volatility ($\sigma^*_{1,t}$),

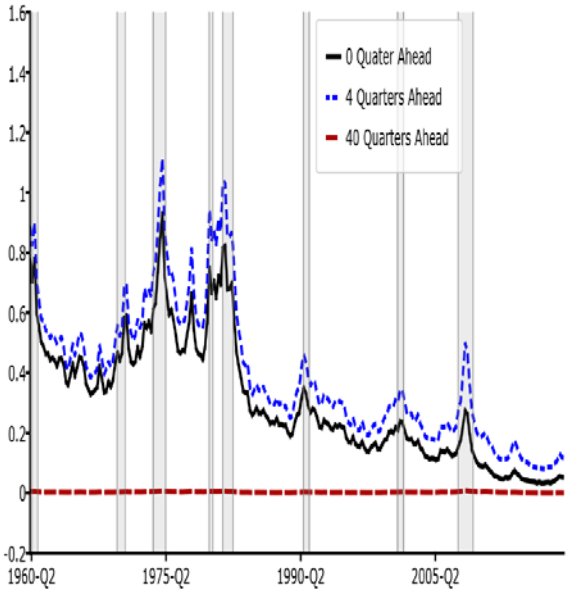


Ratio (r_t)

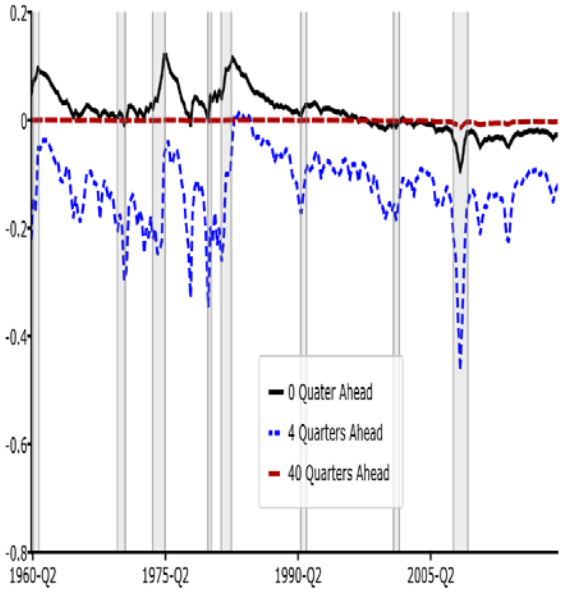
Figure 3. Impulse-Responses [Posterior Medians]



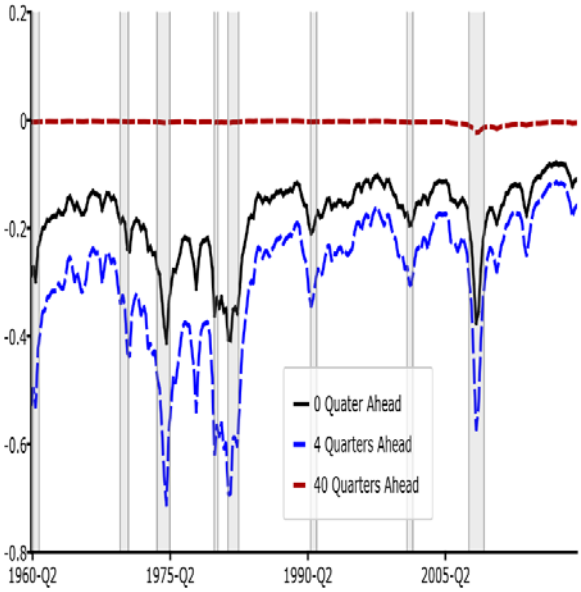
Output Response to Permanent Shock



Output Response to Transitory Shock

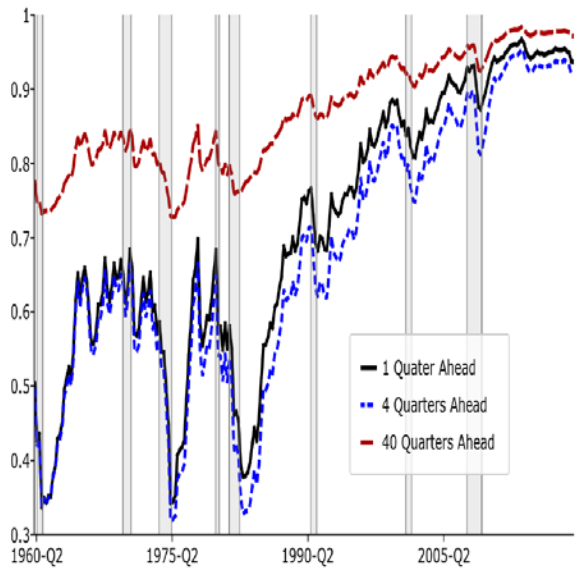


Unemployment Response to Permanent Shock

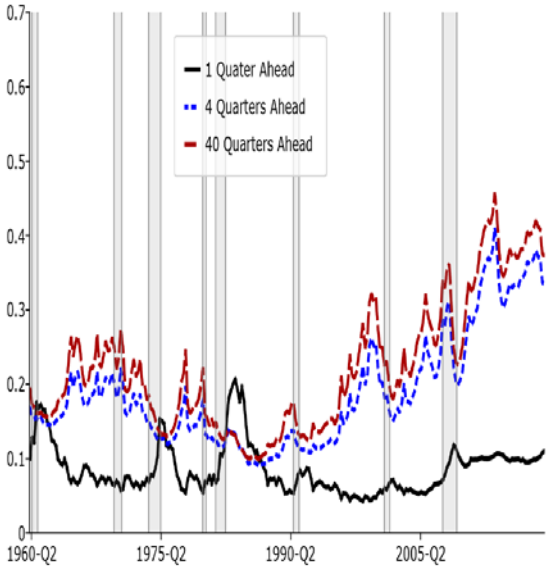


Unemployment Response to Transitory Shock

Figure 4. Variance Decompositions: Variances Due to the Permanent Shock [Posterior Medians]



Real Output



Unemployment Rate

Figure 5. The $a_{12,t}$ Parameter [Posterior Median and 64% Posterior Bands]

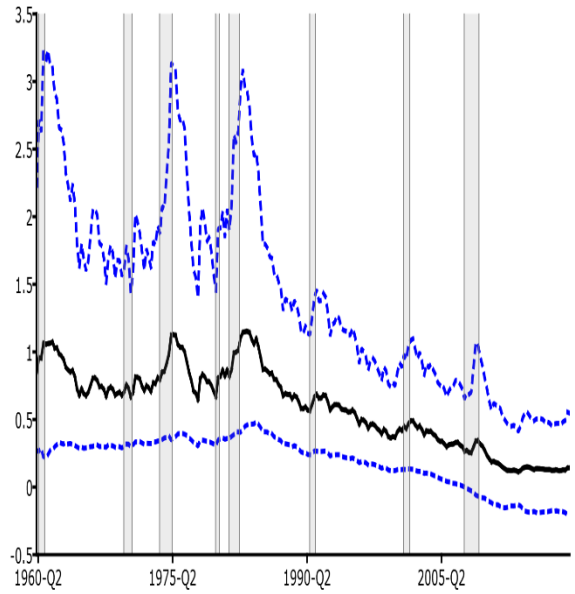
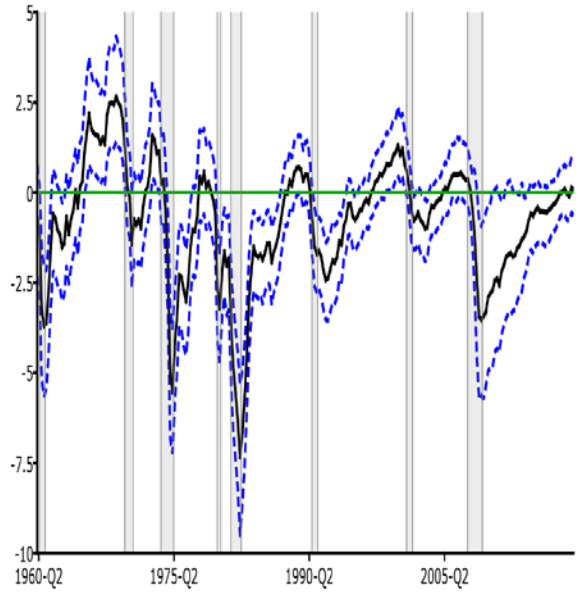
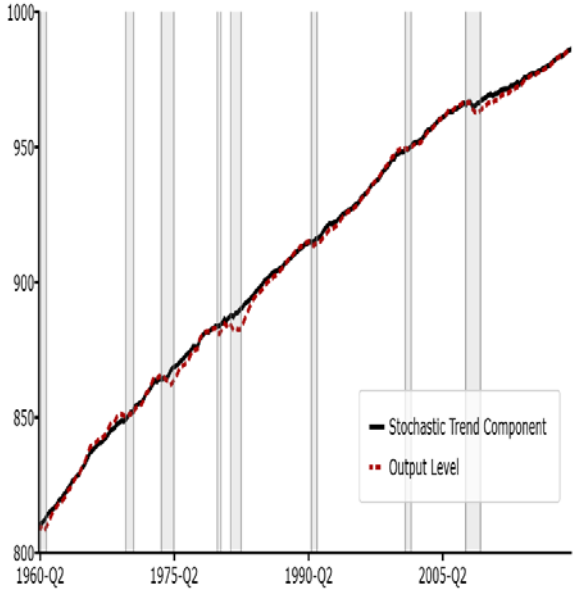


Figure 6.A. Trend-Cycle Decomposition of Real Output $\left(k_{v,a} = 0.03,\ k_{v,b} = 0.04\ \right)$

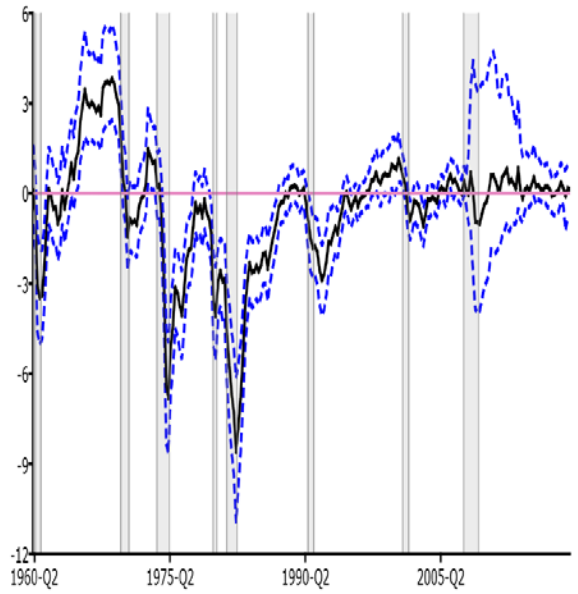


Cycle (Posterior Median, 64% Posterior Bands)

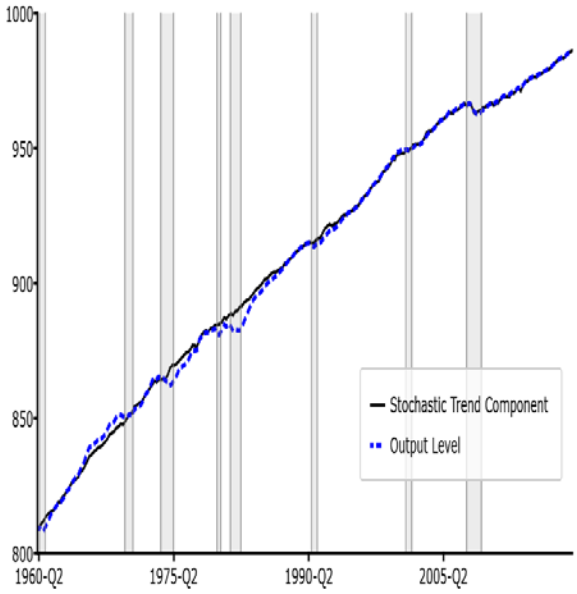


Trend (Posterior Median)

Figure 6.B. Trend-Cycle Decomposition of Real Output $\left(k_{v,a} = 0.03,\ k_{v,b} = 0.03\ \right)$



Cycle (Posterior Median, 64% Posterior Bands)



Trend (Posterior Median)