



Australian School of Business

Working Paper

Never Stand Still

Australian School of Business

Australian School of Business Research Paper No. 2013 ECON15

Structural Evolution of the Postwar U.S. Economy

Yuelin Liu
James Morley

This paper can be downloaded without charge from
The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=2277334>

Structural Evolution of the Postwar U.S. Economy

Yuelin Liu*

University of New South Wales

James Morley

University of New South Wales

June 4, 2013

Abstract

We consider a time-varying parameter vector autoregressive model with stochastic volatility and mixture innovations to study the empirical relevance of the Lucas critique for the postwar U.S. economy. The model allows blocks of parameters to change at endogenously-estimated points of time. Contrary to the Lucas critique, there are large changes at certain points of time in the parameters associated with monetary policy that do not correspond to changes in “reduced-form” parameters for inflation or the unemployment rate. However, the structure of the U.S. economy has evolved considerably over the postwar period, with an apparent reduction in the late 1980s in the impact of monetary policy shocks on inflation, though not on the unemployment rate. Related, we find changes in the Phillips Curve trade-off between inflation and cyclical unemployment (measured as the deviation from the time-varying steady-state unemployment rate implied by the model) in the 1970s and especially since the mid-1990s.

JEL Classification: C11, E24, E32.

Keywords: Time-varying parameters, Mixture innovations, Lucas critique, Great Moderation, Natural Rate of Unemployment, Phillips Curve.

*Corresponding author: Yuelin Liu. Email addresses: yuelin.liu@unsw.edu.au (Yuelin Liu); james.morley@unsw.edu.au (James Morley). We thank Barbara Rossi, Tara Sinclair and participants in the 18th Australasian Macroeconomics Workshop, especially our discussant Joshua Chan, for helpful comments and suggestions. The usual disclaimers apply.

1 Introduction

“[T]he question of whether a particular model is structural is an empirical, not theoretical, one.” – Lucas and Sargent (1981)

The U.S. economy has experienced large shifts in monetary policy regimes since World War II, as discussed by Lucas (1976) and Sargent (1999), amongst many others. Therefore, econometric models designed to study this phenomenon should allow for time-varying parameters. In addition, a substantial decline in the volatility of exogenous shocks, often referred to as “the Great Moderation”, has occurred since mid-1980s. Together, these structural changes imply that a conventional time-invariant vector autoregressive (VAR) model is inadequate for the postwar U.S. data.

In order to integrate time-varying parameters and stochastic volatility of exogenous shocks into an econometric framework, most of the literature has focused on two different approaches: Markov-switching (MS) models and time-varying parameter (TVP) models. MS-VAR models assume that the economy switches between a few (possibly recurrent) regimes abruptly and the magnitude of change across regimes can be large (see, for example, Sims and Zha, 2006). By contrast, TVP-VAR models assume gradual changes (every period of time corresponds to a distinct regime) in conditional mean parameters and the variance-covariance matrix of exogenous shocks (see, for example, Cogley and Sargent (2001, 2005), Primiceri (2005), and Cogley et al. (2010)).

Koop et al. (2009) bridge the two approaches in the literature by extending the stochastic volatility TVP-VAR model proposed by Primiceri (2005) to a framework with mixture innovations. They do so by introducing independent binary latent variables $K_t = (k_{1t}, k_{2t}, k_{3t}), t = 1, 2, \dots, n$ for each block of parameters (3 blocks: conditional mean parameters (k_{1t}), variances (k_{2t}) and covariances (k_{3t}) of exogenous shocks) that allow data to determine the occurrence of break in each block independently in each period of time, such that $Prob(k_{jt} = 1) = p_j$ and $Prob(k_{jt} = 0) = 1 - p_j, j = 1, 2, 3, \forall t$ and corresponding hierarchical priors on p_j , where $p_j \sim Beta(\lambda_{1j}, \lambda_{2j})$. This framework forces the conditional mean parameters to change or stay at previous values all together by construction.¹ However, principal component analysis of the variance-covariance matrix governing the magnitude of shifts in the TVP-VAR parameters conducted by Cogley and Sargent (2005), and also suggested by Sargent (1999), implies that conditional mean parameters change in a highly structured way. Specifically, they argue that the structure of these changes could be due

¹This framework nests the MS-VAR and TVP-VAR models through different prior beliefs on p_j . A pair of small λ_{1j} and large λ_{2j} *a priori* implies infrequent changes, as in the MS-VAR model, while large λ_{1j} and small λ_{2j} implies frequent changes, as in the TVP-VAR model.

to cross-equation restrictions associated with a representative agent’s optimization problem, although imposing these cross-equation restrictions on estimated reduced-form parameters is far from straightforward. Meanwhile, if some reduced-form parameters vary more frequently and more strongly, whilst others are approximately time-invariant, then estimation of a standard TVP-VAR model will tend to overstate variation in some parameters and understate variation in others, which could distort our understanding of the structural evolution of the U.S. economy.

In this paper, we extend Koop et al.’s (2009) analysis to allow for structural changes at different times in subgroups of the VAR parameters, including different blocks of the conditional mean parameters. Because changes in block j of parameters is controlled by a Bernoulli distributed latent variable k_{jt} with $Prob(k_{jt} = 1) = p_j$, the posterior density of p_j reflects the frequency of occurrence of breaks in block j . Then, if the true model is the stochastic volatility TVP-VAR model, as in Primiceri (2005), the data will push the p_j ’s to 1. Otherwise, if the true model is a MS-VAR model, the p_j ’s will be much smaller than 1, with differences in the p_j ’s suggesting different economic forces driving the structural changes. This approach is related to Inoue and Rossi (2011), who allow for a structural break at an unknown break date in subgroups of VAR parameters. However, our model is more flexible in that it allows for multiple shifts in different blocks of parameters.

Building on Koop et al.’s (2009) modeling strategy, our paper makes three contributions: First, because we explicitly divide the VAR parameters into “policy” and “non-policy” blocks, the frequency of changes in “non-policy” blocks relative to that of policy block can be used to test the empirical relevance of Lucas (1976) critique, which states that a shift in systematic policy should induce a change in the “reduced-form” parameters describing the correlations between all of the macroeconomic variables affected by policy. This is different than simulation-based approaches to testing the Lucas critique often considered in the literature; see, for example, Estrella and Fuhrer (2003), Linde (2001), Rudebusch (2005) and Lubik and Surico (2010). Our approach reveals the extent to which the Lucas critique is empirically relevant for the time-varying VAR parameters, including the variances of VAR error terms. We find that Lucas critique is often not relevant. Second, based on standard short-run restrictions, we identify monetary policy shocks and study their effects on inflation and unemployment over time. Our findings can be compared with those in Primiceri (2005) and Koop et al. (2009), who find that there is no statistically significant shift in the effects of monetary policy after World War II, and Kuttner and Mosser (2002) and Boivin and Gianoni (2006), who find that the effects of monetary policy on the U.S. economy has weakened since 1980s. Based on our model, we find that the effects of monetary policy on inflation have only changed at the 3-9 quarter horizon, while the effects of the unemployment rate

have not changed. Third, we estimate the natural rate of unemployment as the time-varying steady-state of the unemployment rate, as in Phelps (1994) and King and Morley (2007). Based on the estimated natural rate, we test for the existence of a Phillips curve trade-off between inflation and cyclical unemployment. We find strong evidence supporting the existence of this short-run trade-off. However, the short-run trade-off has weakened since the late 1970s and has even disappeared since the mid-1990s.

The rest of this paper is organized as follows. Section 2 presents our model. Section 3 describes the data and elicitation of priors. Section 4 considers model comparisons and robustness analysis. Section 5 presents the test for the empirical relevance of Lucas critique. Section 6 reports the results on the evolution of impulse response functions for a monetary policy shock on inflation and the unemployment rate. Section 7 considers the natural rate of unemployment and the short-run trade-off between inflation and cyclical unemployment. Section 8 concludes.

2 Model

One of the contributions of this paper is to allow for structural variations in blocks of VAR parameters by linking them to latent variables K_t 's in the manner of Koop et al. (2009). As a byproduct, we propose a new approach to testing the empirical relevance of Lucas critique that does not rely on simulations from a dynamic stochastic general equilibrium (DSGE) model. We study two intuitive and plausible, although informal and atheoretic, ways of imposing structural changes in reduced-form VAR parameters: (1) by equations; (2) by variables. The details of the model structure are given in the next two subsections.

2.1 A Stochastic Volatility TVP-VAR Model and Identification of a Monetary Policy Shock

The reduced-form TVP-VAR of order p can be cast in the following form:

$$y_t = X_t' \theta_t + \mu_t, \quad \mu_t \sim iid. N(0, \Omega_t)$$

$$X_t = I_n \otimes [1, y_{t-1}', \dots, y_{t-p}'] ,$$

where “ \otimes ” denotes the Kronecker product, y_t is an $n \times 1$ vector including the current observations of endogenous variables, X_t is an $m \times n$ matrix including intercepts and lagged variables, θ_t stacks time-varying reduced-form VAR coefficients and Ω_t is the time-varying variance-covariance matrix of the error term μ_t . In this paper, y_t includes inflation, the

unemployment rate and a short-term interest rate, so $n = 3$ and $m = 21$ because we set $p = 2$ to keep the dimension of parameter space manageable and be consistent with much of the existing literature.²

To identify the monetary policy shock, a structural VAR representation is recovered based on a triangular identification scheme—i.e., we place endogenous variables in the order of $y_t = [\pi_t \ u_t \ i_t]'$, where π_t, u_t, i_t are inflation, the unemployment rate and the short-term interest rate, respectively. This order of endogenous variables assumes that inflation and unemployment respond to monetary policy shock with at least one-period lag.³ Structural shocks $\varepsilon_t, t = 1, 2, \dots, T$, are identified through Cholesky decomposition of the variance-covariance matrix of the reduced-form error terms as follows:

$$A_t \Omega_t A_t' = \Sigma_t \Sigma_t', \quad A_t^{-1} \Sigma_t \varepsilon_t = \mu_t, \quad \varepsilon_t = \Sigma_t \varepsilon_t, \quad \varepsilon_t \sim iid.N(0, I_3)$$

$$A_t = \begin{bmatrix} 1 & 0 & 0 \\ a_{21,t} & 1 & 0 \\ a_{31,t} & a_{32,t} & 1 \end{bmatrix}_{3 \times 3}, \quad \Sigma_t = \begin{bmatrix} \sigma_{11,t} & 0 & 0 \\ 0 & \sigma_{22,t} & 0 \\ 0 & 0 & \sigma_{33,t} \end{bmatrix}_{3 \times 3}$$

Note that $\varepsilon_t = [\varepsilon_{\pi t} \ \varepsilon_{ut} \ \varepsilon_{it}]'$, where the three elements represent fundamental shocks to inflation, unemployment, and monetary policy, respectively. Then, the reduced-form time-varying VAR model can be rewritten as

$$y_t = X_t' \theta_t + A_t^{-1} \varepsilon_t, \quad \varepsilon_t \sim iid.N(0, \Sigma_t \Sigma_t')$$

$$X_t = I_3 \otimes [1 \ y'_{t-1} \ y'_{t-2}], \quad (1)$$

2.2 Mixture Innovations for Time-Varying Parameters and the Variance-Covariance Matrix

The law of motion for the time-varying parameters θ_t is a driftless random walk, following much of the literature on TVP-VAR models, but with more flexible mixture innovations:

$$\theta_t = \theta_{t-1} + K_t \xi_t, \quad \xi_t \sim iid. N(0, Q), \quad (2)$$

²A trivariate VAR model like ours is quite common in the literature; see, for example, Rotemberg and Woodford (1997), Cogley and Sargent (2001, 2005), Primiceri (2005) and Koop et al. (2009).

³It is well recognized that the order of variables in the recursive identification scheme matters given correlation between the reduced-form errors. However, our results for a monetary policy shock are robust to swapping the order of inflation and the unemployment rate.

where Q is positive definite and K_t is a diagonal matrix whose diagonal elements are latent variables k_{it} , $i = 1, 2, 3, 4$, taking on the value of 1 if a change in the corresponding coefficients and 0 otherwise. We consider two types of restrictions on θ_t . In the first case, slopes in the same equation move together, while in the second case slopes on the same variables move together.⁴ In both cases, the intercepts vary together to capture any changes in the long-run levels of inflation, the unemployment rate, and the nominal interest rate. The controlling matrix K_t in the two cases is denoted by $K_t^{(1)} = \text{diag}\{K_{1t}^{(1)}, K_{2t}^{(1)}, K_{3t}^{(1)}, K_{4t}^{(1)}, K_{2t}^{(1)}, K_{3t}^{(1)}, K_{4t}^{(1)}\}$ (by equations) and $K_t^{(2)} = \text{diag}\{K_{1t}^{(2)}, K_{2t}^{(2)}, K_{2t}^{(2)}, K_{2t}^{(2)}, K_{2t}^{(2)}, K_{2t}^{(2)}, K_{2t}^{(2)}\}$ (by variables), respectively, where

$$K_{1t}^{(1)} = K_{1t}^{(2)} = \begin{bmatrix} k_{1t} & 0 & 0 \\ 0 & k_{1t} & 0 \\ 0 & 0 & k_{1t} \end{bmatrix}, \quad K_{2t}^{(1)} = \begin{bmatrix} k_{2t} & 0 & 0 \\ 0 & k_{2t} & 0 \\ 0 & 0 & k_{2t} \end{bmatrix}, \quad K_{3t}^{(1)} = \begin{bmatrix} k_{3t} & 0 & 0 \\ 0 & k_{3t} & 0 \\ 0 & 0 & k_{3t} \end{bmatrix},$$

$$K_{4t}^{(1)} = \begin{bmatrix} k_{4t} & 0 & 0 \\ 0 & k_{4t} & 0 \\ 0 & 0 & k_{4t} \end{bmatrix}, \quad K_{2t}^{(2)} = \begin{bmatrix} k_{2t} & 0 & 0 \\ 0 & k_{3t} & 0 \\ 0 & 0 & k_{4t} \end{bmatrix}.$$

In terms of the variance-covariance matrix for the VAR errors, let α_t be a vector collecting the non-diagonal and non-zero elements in A_t and σ_t be a vector collecting the diagonal elements in Σ_t . Then the evolution of elements in α_t and σ_t is as follows:

$$\alpha_t = \alpha_{t-1} + k_{5t}\eta_t, \quad \eta_t \sim iid. N(0, S), \quad (3)$$

$$\ln \sigma_t = \ln \sigma_{t-1} + k_{6t}\zeta_t, \quad \zeta_t \sim iid. N(0, W), \quad (4)$$

where S, W are positive definite and S is block diagonal in which each block is corresponding to parameters in different equations and similarly, $k_{jt} = 1$, $j = 5, 6$, if a change in the subset of parameters occurs and $k_{jt} = 0$, $j = 5, 6$, otherwise.

We assume that all of the innovation blocks in the dynamic system are uncorrelated contemporaneously and at all lags and leads—i.e., they are jointly normally distributed with

⁴Cogley and Sargent (2005) conduct a simple principal components analysis and find that the parameters in the inflation equation receive substantially higher loadings compared to the rest of the parameters—i.e., the variations in the reduced-form VAR parameters are highly structured. As demonstrated in Sargent (1999), the structured variation can be derived from the cross-equation restrictions associated with an agent's optimization problem. However, it is not so clear-cut how to impose the relevant cross-equation restrictions in practice. Thus, from an empirical perspective, we investigate different ways of allowing subsets of parameters to change at the same time and let the data speak as to which categorization best fits the historical experience.

the following variance-covariance matrix V as:

$$V = Var \left(\begin{bmatrix} \epsilon_t \\ \xi_t \\ \eta_t \\ \zeta_t \end{bmatrix} \right) = \begin{bmatrix} I_3 & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix}.$$

Estimation of parameters in this framework relies on Markov Chain Monte Carlo (MCMC) methods. In particular, under the assumption that the k_{jt} 's are independent of one another, contemporaneously and at all lags and leads, K_t, k_{5t}, k_{6t} can be drawn based on the reduced conditional sampling algorithm proposed in Gerlach et al. (2000), without conditioning on the state vector θ_t . This greatly improves the efficiency of the sampler when K_t and θ_t are highly correlated.⁵ Then, following Primiceri (2005), we adapt methods in Carter and Kohn (1994) and Kim et al. (1998) to draw state vectors θ_t , α_t and $\ln \sigma_t$ from three Gaussian linear state-space systems separately. See the detailed MCMC procedures in the technical appendix.

3 Data and Priors

3.1 Data

As discussed above, we use a small three-variable VAR to study the evolution of U.S. monetary policy, measured by short-term nominal interest rate (federal funds rate, averaged from daily rates, series ID: FEDFUNDS), and its impacts on inflation (seasonally adjusted compounded annual rate of change of Personal Consumption Expenditures, series ID: PCEPTPI) and the unemployment rate (seasonally adjusted civilian unemployment rate, all workers over 16, series ID: UNRATE).⁶ The series are quarterly and run from 1954:Q3 to 2007:Q4.

⁵Even though it might be desirable to explicitly model correlations between the k_{jt} 's, this would bring a huge cost in terms of increasing the computation burden, making implementation extremely difficult. In any event, any dependence in the timing of changes across different parameter blocks should be evident from *ex post* correlations in the posterior estimates of changes in the different parameters.

⁶All of the data series were downloaded from FRED managed by Federal Reserve Bank of St. Louis at <http://research.stlouisfed.org/fred2/>. The results are robust to different measurements of inflation and short-term interest rates, for example, GDP deflator and 3-month treasury bill rate. Personal Consumption Expenditures (PCE) is used because it has been, at least in recent years, the Fed's preferred measure of the cost of living. When 3-month treasury bill rate is employed instead of federal funds rate, an unemployment puzzle (i.e., a contraction monetary policy associated with a decline in unemployment rate) appears, as in Primiceri (2005) and Koop et al. (2009). Therefore, we take the federal funds rate (FFR) which is directly under control of the Fed as the monetary policy instrument. Another potential concern is that the FFR might be less representative as a policy instrument from 1979 to 1982 during which the Fed was officially framed its policy in terms of monetary aggregates. But Cook (1989) argues that even in that episode the

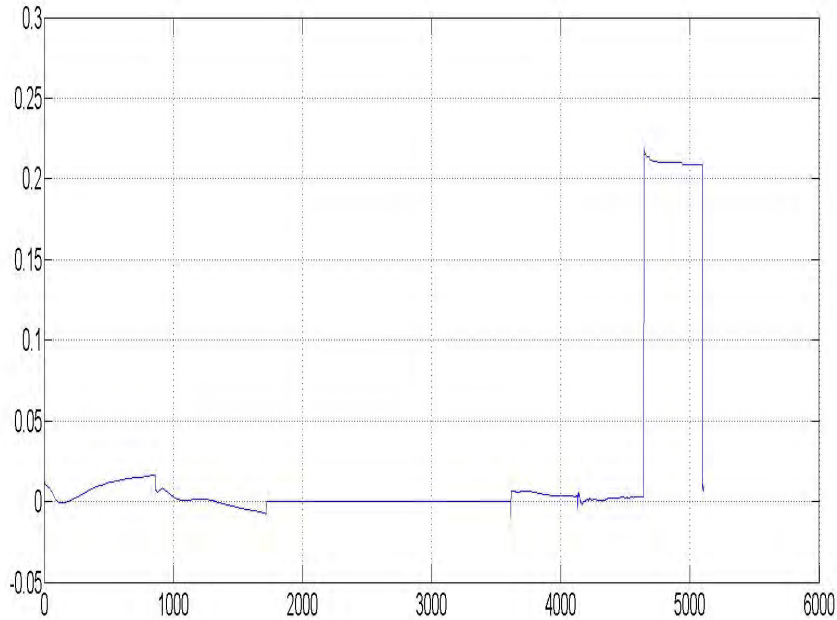


Figure 1: Sample 20th-order Autocorrelation Functions for All of the Parameters

We execute 70,000 replications of the Gibbs sampler, with the first 20,000 draws, known as the “burn-in”, discarded to allow for convergence to the ergodic distribution. Every 10th draw is saved from the remaining 50,000 draws to economize the storage space. Therefore, Bayesian inferences are carried out based on 5,000 draws from the posterior distribution. Convergence diagnostics are conducted by inspecting sample ACFs and recursive means of all of the parameters. As shown in Figure 1, sample 20th-order ACFs for all of the parameters (including hyperparameters) are well below or slightly over 0.2 implying that the posterior draws mix well and the convergence check is satisfactory.

3.2 Priors

Priors for state vectors and hyperparameters are calibrated following Primiceri (2005) and Koop et al. (2009), with a few minor modifications. Data for the first ten years of the sample (42 observations, 1954:Q3 – 1964:Q4) are employed to calibrate the priors. Specifically, a time-invariant VAR model is estimated using conditional MLE, which produces point estimates for VAR coefficients, $\hat{\theta}_0$ and its corresponding variances $V(\hat{\theta}_0)$. Estimates of the variance-covariance matrix of VAR error terms, $\hat{\Omega}_0$, are obtained as well and $\hat{\alpha}_0$, $\hat{\sigma}_0$ is derived from decomposing $\hat{\Omega}_0$. The variance of $\hat{\alpha}_0$, $V(\hat{\alpha}_0)$, is obtained by simulation from a Wishart

FFR serves as a satisfactory policy indicator. Hence, it seems appropriate to treat the FFR as the policy instrument across the whole sample.

distribution with scatter matrix $\hat{\Omega}_0$ and degree of freedom = 40, and we set the variance of $\ln(\hat{\sigma}_0)$ to $10I_3$ which is large in log-scale, implying a small weight is put on the prior. As for the hyperparameters Q , $S = \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}$, W , the priors are inverse-Wishart distributions. In order to put as least weight as possible on prior beliefs, the degree of freedom corresponding to each inverse-Wishart distribution is set to the minimum plausible value $\dim(Q) + 1 = 22$, $\dim(S_1) + 1 = 2$, $\dim(S_2) + 1 = 3$, $\dim(W) + 1 = 4$, respectively. In summary, the priors are as follows:

$$\begin{aligned}\theta_0 &\sim N(\hat{\theta}_0, 4V(\hat{\theta}_0)), \\ \alpha_0 &\sim N(\hat{\alpha}_0, 4V(\hat{\alpha}_0)), \\ \ln \sigma_0 &\sim N(\ln \hat{\sigma}_0, 10I_3), \\ Q &\sim IW(40k_Q^2 V(\hat{\theta}_0), 22), \\ S_1 &\sim IW(2k_S^2 V(\hat{\alpha}_{1,0}), 2), \\ S_2 &\sim IW(3k_S^2 V(\hat{\alpha}_{2,0}), 3), \\ W &\sim IW(4k_Q^2 I_3, 4),\end{aligned}$$

where $k_Q = k_W = 0.01$, $k_S = 0.1$ and $\hat{\alpha}_{1,0}, \hat{\alpha}_{2,0}$ correspond to each block of $\hat{\alpha}_0$.⁷

To complete the model, hierarchical priors for K_t , k_{5t} and k_{6t} need to be specified. We adopt a Bernoulli distribution $Ber(p_j)$ with

$$Prob(k_{jt} = 1) = p_j, \quad j = 1, 2, \dots, 6, \quad (5)$$

where p_j is the probability of a parameter change occurring at time t for k_{jt} . The prior for p_j is a Beta distribution $Beta(\lambda_{1j}, \lambda_{2j})$, $j = 1, 2, \dots, 6$, which forms a conjugate prior with a Bernoulli distribution. The values of λ_{1j} and λ_{2j} reflect prior beliefs about the frequency of parameter changes in the model. Small values of λ_{1j} and large values of λ_{2j} imply a “structural break” (SB) model with few changes in the parameters (e.g., $\lambda_{1j} = 0.01$, $\lambda_{2j} = 10$ for all j , would imply $E(p_j) = 0.001$, $sd(p_j) = 0.01$). Large values of λ_{1j} and small values of λ_{2j} approaches the standard stochastic volatility TVP model in Primiceri (2005) (e.g., $\lambda_{1j} = 1$, $\lambda_{2j} = 0.01$ for all j , would imply $E(p_j) = 0.99$, $sd(p_j) = 0.08$). For our benchmark model, we set $\lambda_{1j} = \lambda_{2j} = 1$ for all j , such that $E(p_j) = 0.5$, $sd(p_j) = 0.29$, meaning that *a priori* we believe the occurrence of a change in each period occurs with a 50% chance. We also consider SB and TVP versions of the models in which priors for the p_j 's are extremely tight, as suggested above. In general, we combine these priors with those for state vectors

⁷See Primiceri (2005) for a full discussion of the reasons behind these values of k_Q, k_S, k_W .

$\theta_t, \alpha_t, \sigma_t$ and hyperparameters Q, S, W . Meanwhile, Bayesian inferences in the benchmark models with $K_t^{(1)}$ or $K_t^{(2)}$ rely on diffuse priors that let data speak as much as possible. Comparisons of model fit are investigated in the next section.

4 Model Fit and Robustness Analysis

Table 1 reports on the fit of the benchmark model, the SB model, and the TVP model using posterior means of p_j 's and the expected value of log-likelihood function as described in Carlin and Louis (2000), rather than marginal likelihoods that are sensitive to priors and difficult to calculate, especially in this setting of models with extremely high-dimensional parameter space. The label BEQ stands for the benchmark model varying with respect to equations, whereas BVA denotes the benchmark model varying with respect to variables. SBEQ, SBVA, TVPEQ and TVPVA are analogous labels for the versions of the model with strong priors on the p_j 's.

The first result that stands out in Table 1 is that the BEQ model with structural variations in intercepts and slopes in different equations receives strongest support based on the expected log-likelihood $E(\log L|Y)$.⁸ In accordance with Primiceri (2005), Cogley and Sargent (2005), and Koop et al. (2009), the probabilities of parameter change for the BEQ model suggest that the volatility of error terms has changed frequently over time with $E(p_6|Y) = 0.9713$ (standard deviation 0.0241). However, the slopes in unemployment equation are relatively stable with $E(p_3|Y) = 0.2844$ (standard deviation 0.0546), which is substantially smaller than the probabilities of parameter change in other blocks. This implies that, on average, a change occurs in the slopes in the unemployment equation every 4 quarters, whereas intercepts and slopes in inflation and interest rate equations are expected to change every one or two quarters. These results provide strong evidence for the idea that changes in reduced-form VAR parameters are highly structured.

Second, we find that, even when extremely tight priors on p_j 's are considered with the SB and TVP versions of the model, the information in the data is so strong that it pushes posterior inferences much of the distance towards the results for the benchmark model. This can be treated as robustness analysis for our modeling strategy. For example, in the SBEQ model, the parameters of $Beta(\lambda_{1j}, \lambda_{2j})$ priors on p_j 's are set to $\lambda_{1j} = 0.01, \lambda_{2j} = 10, j = 1, 2, \dots, 6$, and $E(p_j) = 0.001, sd(p_j) = 0.01$ which means that on average a break is expected to happen once every 1000 quarters *a priori*. Nevertheless, the posterior mean values of p_j 's show

⁸The measure of the expected log-likelihood is obtained by averaging the log-likelihood from the state-space model (1) and (2) based on each draw of $\alpha^T, \sigma^T, Q, S, W, K^T, k_5^T, k_6^T, \lambda$, where (and hereafter) $x^T = (x_1, x_2, \dots, x_T)$.

Table 1: Model comparison

Models	$E(p_1 Y)$	$E(p_2 Y)$	$E(p_3 Y)$	$E(p_4 Y)$	$E(p_5 Y)$	$E(p_6 Y)$	$E(\log L Y)$
BEQ	0.9386 (0.0278)*	0.9026 (0.0330)	0.2844 (0.0546)	0.4866 (0.0588)	0.6717 (0.2295)	0.9713 (0.0241)	3733.5
BVA	0.8162 (0.0382)	0.8604 (0.0319)	0.4494 (0.0604)	0.6686 (0.0434)	0.6755 (0.2392)	0.9700 (0.0247)	1877.6
TVPEQ	0.9990 (0.0068)	0.9992 (0.0079)	0.5041 (0.1862)	0.9979 (0.0173)	0.9991 (0.0076)	0.9997 (0.0025)	1261.1
TVPVA	0.9996 (0.0052)	0.9993 (0.0070)	0.7487 (0.2281)	0.9948 (0.0244)	0.9986 (0.0125)	0.9997 (0.0023)	1334.2
SBEQ	0.6847 (0.0481)	0.5998 (0.0605)	0.3951 (0.0487)	0.5105 (0.0473)	0.0955 (0.0870)	0.8370 (0.0417)	2275.4
SBVA	0.6984 (0.0392)	0.5611 (0.0406)	0.3496 (0.0387)	0.8010 (0.0345)	0.0962 (0.0875)	0.7996 (0.0486)	2190.7

*Standard deviations are listed in parentheses.

that $E(p_1|Y) = 0.6847$, $E(p_2|Y) = 0.5998$, $E(p_3|Y) = 0.3951$, $E(p_4|Y) = 0.5105$, $E(p_5|Y) = 0.0955$, $E(p_6|Y) = 0.8370$ with standard deviations 0.0481, 0.0605, 0.0487, 0.0473, 0.0870 and 0.0417, respectively. Except for the block of contemporaneous cross-equation impacts of structural shocks, α_t , the posterior expected values of p_j 's suggest that the parameter blocks are changing approximately every one or two quarters, which strongly rejects our prior belief that few breaks occur over time. On the other hand, in the TVPEQ model, λ_{1j} and λ_{2j} are set to 1 and 0.01, respectively. So *a priori* $E(p_j) = 0.99$ with standard deviation 0.08 favors a time-varying parameter with stochastic volatility model. However, the posterior mean values of p_1 up to p_6 are 0.9990, 0.9992, 0.5041, 0.9979, 0.9991 and 0.9997, respectively. The probability of observing a break in the slopes in unemployment equation at every time period declines substantially from 0.9900 to 0.5041 with standard deviation 0.1862 which significantly deviates from the prior belief. This suggests that the slopes in the unemployment rate equation is more stable than the other blocks in the model and supports the idea of highly structured changes in the reduced-form VAR parameters.

In summary, our BEQ model receives the strongest support from the data. Hereafter, the empirical results about the evolution of US economy are based on this model.

5 Testing the Lucas Critique

Ever since Lucas's (1976) seminal paper, it has been widely-recognized by macroeconomists that reduced-form econometric models could be inappropriate for policy analysis if there are

changes in parameters describing policy. However, a relatively large literature—see Fevero and Hendry (1992), Estrella and Fuhrer (2003), Linde (2001) and Rudebusch (2005), among others—casts doubt on the empirical relevance of Lucas critique by considering Chow tests and superexogeneity tests. In a recent paper, Lubik and Surico (2010) find that, by taking stochastic volatility in the reduced-form errors into account, one cannot reject the empirical relevance of Lucas critique. Specifically, a shift in policy rule has a great impact not only on reduced-form conditional mean parameters, but also on the variances of reduced-form error terms. They criticize the Chow and superexogeneity tests employed in Fevero and Hendry (1992), Estrella and Fuhrer (2003) and Rudebusch (2005), and others for implicitly assuming homoskedasticity of the reduced-form error terms, undermining the power of the tests.⁹

Existing approaches for testing the Lucas critique typically rely on simulating data from a specified DSGE model as if the model were the “true” data generating process (DGP) of the macro variables of interest. By contrast, we make no assumption that the metaphors involved in a given modeling approach are literally true when testing the Lucas critique. Instead, our approach lets the data speak as to whether “policy” block and “non-policy” block parameters change at the same time.

Ideally, we would like to be able to directly identify changes in the parameters for a structural policy equation. However, our approach only allows us to consider changes in blocks of conditional mean parameters for the reduced-form VAR model.¹⁰ But this is less of a problem than it may at first appear because a shift in parameters for the structural policy equation should induce a shift in the parameters for the reduced-form policy equation. Then, if the Lucas critique holds, the shift in the policy parameters should also induce a shift in the parameters of the reduced-form equations for the non-policy variables, at least for variables affected by monetary policy. It is true that a simultaneous shift in the reduced-form parameters for the policy and non-policy equations could instead be due to a change in non-policy structural parameters, as suggested by Inoue and Rossi (2011). But a shift in

⁹Another potential issue is the low power of the superexogeneity test in small samples, as discussed in Linde (2001) and Collard et al.(2001).

¹⁰Although we need to consider changes in conditional mean parameters for the reduced-form VAR model, our approach does allow for identification of shifts in the structural shock variances and the contemporaneous cross-equation impact of the structural shocks on the observables. Also, it should be noted that we would be able to directly identify changes in the structural policy equation if our identification of monetary policy shocks involved placing the interest rate first rather than last in the causal ordering. This would correspond to the idea that policy only responds to inflation and the unemployment rate with a lag, which could be justified based on data availability issues. We note that the impulse responses for a policy shock for this alternative ordering are qualitatively similar to those for the standard ordering employed in our analysis. Also, the results for the Lucas critique tests are similar for this alternative identification (the fact there are any differences being due to the fact that the ordering matters for the identification of the structural shock variances and the contemporaneous cross-equation impact of the structural shocks on observables, which are linked to k_5 and k_6 , respectively).

Table 2: Posterior median values of fractions of $k_{it} = k_{jt}, i, j = 1, 2, \dots, 6, i \neq j$

$k_{it} = k_{jt}, i \neq j$	k_2	k_3	k_4	k_5	k_6
k_1	0.8605 (0.0286)*	0.2733 (0.0504)	0.5000 (0.0422)	0.6919 (0.2042)	0.9244 (0.0279)
k_2	–	0.3256 (0.0385)	0.5000 (0.0355)	0.6744 (0.1879)	0.8895 (0.0300)
k_3	–	–	0.4942 (0.0343)	0.4128 (0.1093)	0.2849 (0.0429)
k_4	–	–	–	0.5000 (0.0423)	0.4884 (0.0459)
k_5	–	–	–	–	0.7035 (0.2196)

*Standard deviations are listed in parentheses.

the reduced-form policy parameters without a corresponding shift in the reduced-form non-policy parameters should only occur if the Lucas critique does not hold (again, assuming policy impacts the relevant variable). It is this possibility that we consider in our empirical analysis.

Table 2 reports on the timing of breaks across parameter blocks. The reduced-form intercepts (controlled by k_1), reduced-form slopes in the inflation equation (controlled by k_2), reduced-form slopes in the interest rate equation (controlled by k_4), contemporaneous cross-equation impacts of structural shocks (controlled by k_5), and standard deviations of structural shocks (controlled by k_6) co-move frequently with the probability of co-movement varying from 49% to as high as 92%. By contrast, the relationship between the reduced-form slopes in the unemployment equation (controlled by k_3) and other parameter blocks suggests much less pairwise dependence, with co-movement varying between 27% and 49%. It should be noted that this relatively weak pairwise dependence between k_3 and $k_i, i \neq 3$ is not the result of the unemployment rate being unrelated to the interest rate or inflation. As discussed in the next section, monetary policy shocks have significant effects on the unemployment rate throughout the sample period. So, if the Lucas critique holds, the reduced-form parameters should all move together.

As an even more direct way to look at co-movement related to changes in policy parameters, Table 3 reports on shifts in blocks of parameters conditional on a change in slope parameters for the interest rate equation. The results initially suggest support for Lucas critique for the behaviour of inflation and the variance-covariance matrix (consistent with Lubik and Surico, 2010), although, again, not for the behaviour of the unemployment rate. Specifically, given a change in slope parameters for the interest rate equation, there is more

Table 3: Test of the Lucas critique: Posterior probabilities of changes in parameter blocks conditional on shift in slope parameters in the interest rate equation

	Probability	95% Credible Interval
Inf ^a	0.9236 ^b	[0.9175, 0.9315]
Unem	0.2530	[0.2424, 0.2925]
NP	0.9356	[0.9309, 0.9474]
VarErr	0.9924	[0.9920, 0.9951]

(a). Inf: slopes in inflation equation; Unem: slopes in unemployment equation; NP: non-policy block; VarErr: variances of reduced-form error terms. (b). Probability: posterior medians.

than a 90% probability of a change in the other parameters, except for the slopes in the unemployment rate equation, which only have a 25% conditional probability of change. However, looking back at Table 1, it is clear that certain parameters almost always change. For example, the structural shock variances appear to change about 97% of the time. So whenever the interest rate slope parameters change 48% of the time, we would expect a high conditional probability that the variance-covariance matrix changes too. But the question remains as to whether the parameter changes are related to each other.

To answer the question of whether simultaneous parameter changes are coincidental, we calculate the correlation between changes in policy parameters and non-policy parameter blocks conditional on a change in slope parameters for the interest rate equation. Table 4 reports posterior inferences for these correlations. Most of the correlations are essentially zero. That is, even if policy and non-policy parameters change at the same time, they do not change together in a systematic fashion. Thus, the evidence argues against the Lucas critique. Nonetheless, there are some significant correlations, suggesting that the zero correlations are not merely the consequence of the modeling assumption of independent switching in different blocks of parameters. For example, the most statistically significant correlation is $corr(\Delta\bar{\beta}_{it}^i, \Delta\bar{\beta}_{it}^u) = -0.1968$, which suggests that an increase in interest rate smoothing has a weak negative association with the impact of the interest rate on the unemployment rate. Other small, but possibly non-zero correlations suggest that a smaller weight for policy on the unemployment rate corresponds to a smaller impact of the interest rate on the unemployment rate $corr(\Delta\bar{\beta}_{ut}^i, \Delta\bar{\beta}_{it}^u) = -0.1292$ and a larger trade-off between the unemployment rate and inflation $corr(\Delta\bar{\beta}_{ut}^i, \Delta\bar{\beta}_{ut}^\pi) = -0.1922$. Also, a larger weight for policy responses on inflation correspond to a larger trade-off between the unemployment rate and inflation $corr(\Delta\bar{\beta}_{\pi t}^i, \Delta\bar{\beta}_{ut}^\pi) = -0.1777$ and less persistent inflation $corr(\Delta\bar{\beta}_{\pi t}^i, \Delta\bar{\beta}_{\pi t}^\pi) = -0.1146$. But these small correlations are a far cry from the idea inherent in the Lucas critique that changes in the transmission mechanism of monetary policy, changes in the

Table 4: Contemporaneous cross correlations between changes in policy parameters and non-policy parameters conditional on shift in slope parameters in the interest rate equation

Non-Policy Block	Policy Block				
	$\Delta\bar{\beta}_{\pi t}^i$	$\Delta\bar{\beta}_{ut}^i$	$\Delta\bar{\beta}_{it}^i$	$\Delta a_{31,t}$	$\Delta a_{32,t}$
$\Delta\bar{\beta}_{\pi t}^\pi$	-0.1146 (0.1009)*	-0.0652 (0.0984)	0.0612 (0.1001)	0.0007 (0.0766)	-0.0009 (0.0770)
$\Delta\bar{\beta}_{ut}^\pi$	-0.1777 (0.0996)	-0.1922 (0.0977)	-0.0960 (0.1002)	0.0004 (0.0759)	0.0008 (0.0780)
$\Delta\bar{\beta}_{it}^\pi$	-0.0248 (0.0999)	-0.0855 (0.0971)	-0.0132 (0.0979)	-0.0008 (0.0766)	0.0005 (0.0775)
$\Delta\bar{\beta}_{\pi t}^u$	-0.0078 (0.0796)	-0.0268 (0.0786)	-0.0814 (0.0781)	0.0006 (0.0766)	0.0001 (0.0767)
$\Delta\bar{\beta}_{ut}^u$	-0.0162 (0.0787)	-0.0294 (0.0786)	-0.0381 (0.0794)	-0.0012 (0.0759)	0.0005 (0.0762)
$\Delta\bar{\beta}_{it}^u$	-0.0575 (0.0787)	-0.1292 (0.0793)	-0.1968 (0.0801)	0.0010 (0.0766)	0.0001 (0.0761)
$\Delta a_{21,t}$	-0.0003 (0.0774)	0.0018 (0.0780)	0.0033 (0.0781)	0.0007 (0.1051)	0.0022 (0.1058)

*Standard deviations are listed in parentheses. Δ is the difference operator. $\bar{\beta}_{it}^\pi$, $\bar{\beta}_{\pi t}^u$ and $\bar{\beta}_{ut}^i$ are sums of slopes on the interest rate in the inflation equation, slopes on inflation in the unemployment rate equation, and slopes on the unemployment rate in the interest rate equation, respectively. The other $\bar{\beta}$'s are analogous.

slope of the Phillips curve, or changes in the persistence of inflation are driven primarily by changes in the policy regime.

6 Evolution of Impulse Responses

There is considerable academic debate about whether monetary policy is responsible for stabilizing the US economy since the mid-1980s, a phenomenon known as the ‘‘Great Moderation’’. One way to investigate the potential sources of the decline in volatility is to consider changes in the responses of macroeconomic variables to monetary policy shocks over time. Along these lines, Primiceri (2005) and Koop et al. (2009) find no evidence for a major role played by monetary policy because they find that impulse responses for inflation and the unemployment rate to a monetary policy shock do not change significantly over time. However, Kuttner and Mosser (2002) and Boivin and Giannoni (2006) find that the impact of monetary policy on output and inflation appears somewhat weaker in recent years compared

to before 1980s.¹¹

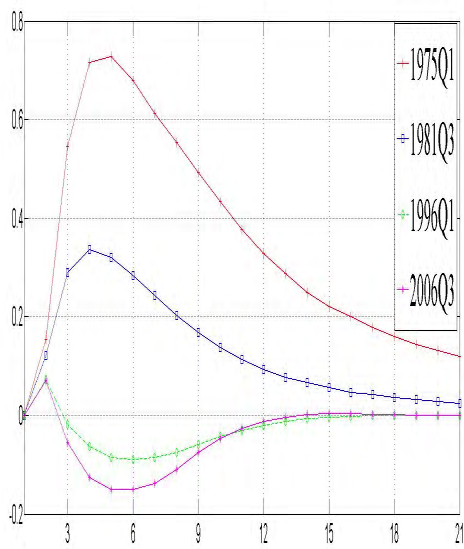
Figures 2-4 plot the evolution of impulse responses for inflation, the unemployment rate, and the interest rate to a 1 percent monetary policy shock at selected dates: 1975Q1, 1981Q3, 1996Q1 and 2006Q3.¹² The estimated magnitude of the responses of inflation, the unemployment rate, and the interest rate is generally smaller since 1980s. However, the differences from the 1975Q1 responses are not statistically significant except for the responses of inflation at the 3-9 quarter horizon.

Our results are at odds with Primiceri (2005), Koop et al. (2009) and Boivin and Giannoni (2006).¹³ Boivin and Giannoni (2006) study time-invariant VAR models of inflation, output, and the interest rate for two subsamples 1959:Q1 - 1979:Q3 and 1979:Q4 - 2002:Q2 and compare impulse responses evaluated from subsamples based on the recursive identification scheme. However, they only show point estimates of the impulse responses without conducting a rigorous statistical test of whether impulse responses have actually changed across subsamples. As for Primiceri (2005) and Koop et al. (2009), their modeling strategy is quite similar to ours. Thus, it is fairly easy to determine the source of the different results. Specifically, every structural parameter is a mapping from the reduced-form VAR parameters given a particular identification scheme. Then, the impulse responses are functions of the structural parameters. Therefore, if the reduced-form VAR parameter estimates are misleading due to model misspecification, the impulse response functions will be contaminated as well. As discussed in Sections 3 and 4, the TVPEQ and TVPVA models with the tight priors implying a break in parameters each period of time is essentially the same as the stochastic volatility TVP model in Primiceri (2005). Also, our benchmark models, BEQ and BVA, would collapse to the model in Koop et al. (2009) if parameters change or stay the same simultaneously. However, as clearly shown in Table 1, the BEQ model receives the strongest

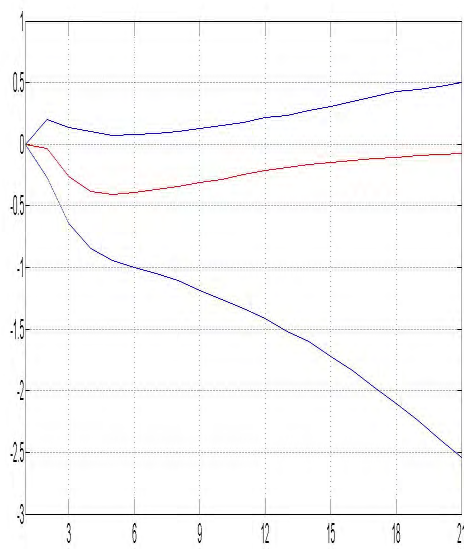
¹¹Another way to investigate sources of the decline in volatility is to consider counterfactual analysis. Using this approach, Sims and Zha (2006) find that smaller shocks are responsible, while Inoue and Rossi (2011) find that a change in monetary policy also played a role.

¹²For easy comparison, these dates are the same as those considered in Primiceri (2005) and Koop et al. (2009). However, we consider the Fed Funds Rate and inflation based on the PCE deflator, while they consider the 3-month treasury bill rate as a proxy for the policy instrument and measure inflation using the GDP deflator. Despite this difference, it should be noted that the impulse response results are robust to considering the other measures of the interest rate and inflation.

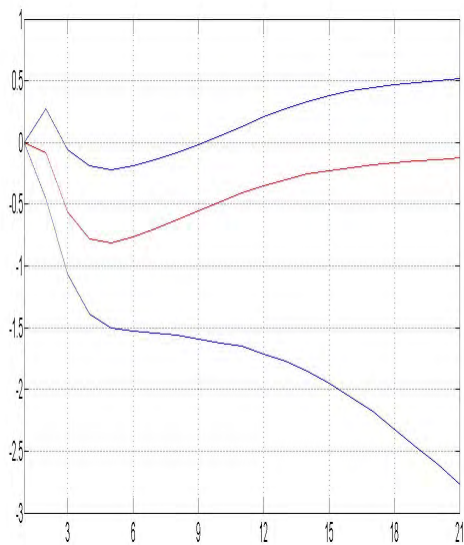
¹³There is an apparent “price puzzle” in 1975Q1 and 1981Q3, which is common for small monetary VAR models with triangular identification schemes for pre-1980 US data. This might suggest misspecification of the model—i.e., some informative variables that impact the Fed and private sectors’ decision-making processes are missing from the model. As suggested by Sims (1992), one promising way to solve this problem is to include a commodity price index. Nevertheless, for the sake of computational feasibility given the already large dimension of the parameter space, we stick with the trivariate model. Also, we are interested in the evolution of impulse responses instead of impulse responses *per se*. Thus, the price puzzle should not be as much of a hindrance for understanding variations in impulse responses as it is for understanding the responses themselves.



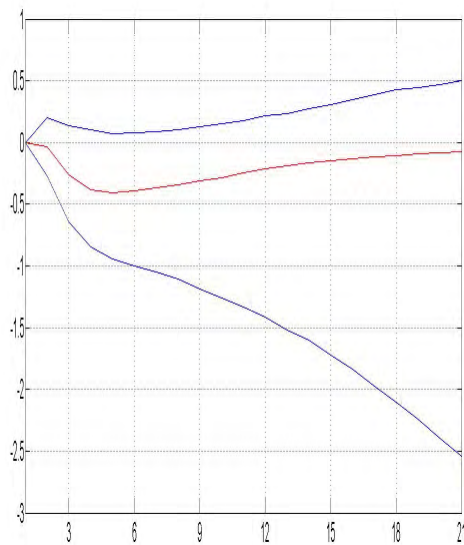
(a)



(b)

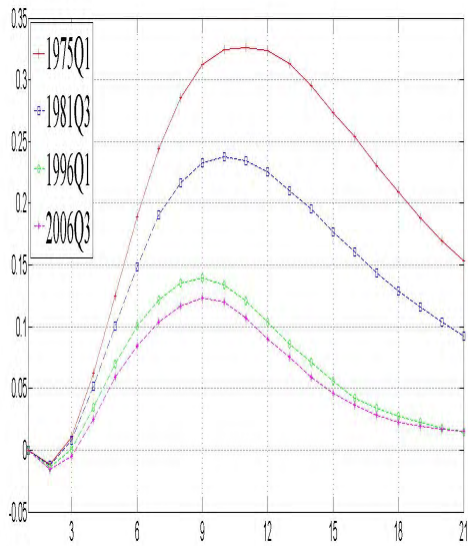


(c)

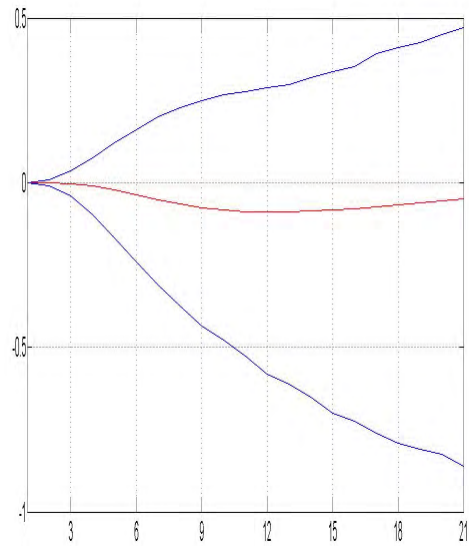


(d)

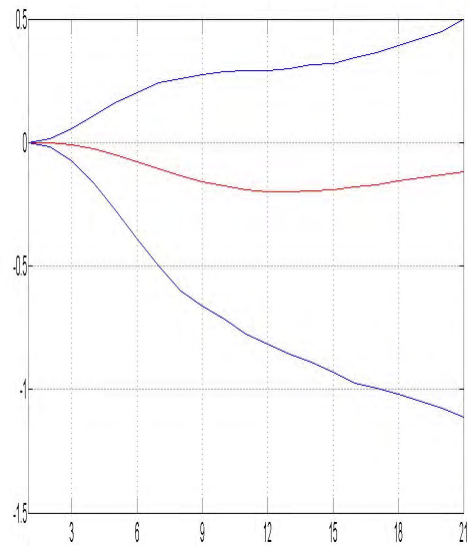
Figure 2: Response of inflation to a 1 percent monetary policy shock in 1975Q1, 1981Q3, 1996Q1 and 2006Q3: (a) medians of impulse responses; (b) response in 1981Q3 minus response in 1975Q1 with 90% credible interval; (c) response in 1996Q1 minus response in 1975Q1 with 90% credible interval; (d) response in 2006Q3 minus response in 1996Q1 with 90% credible interval.



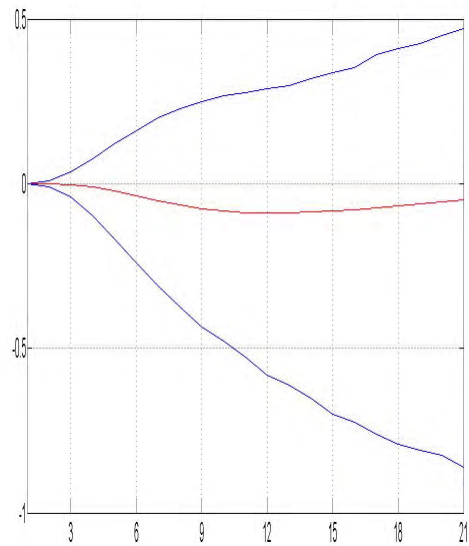
(a)



(b)

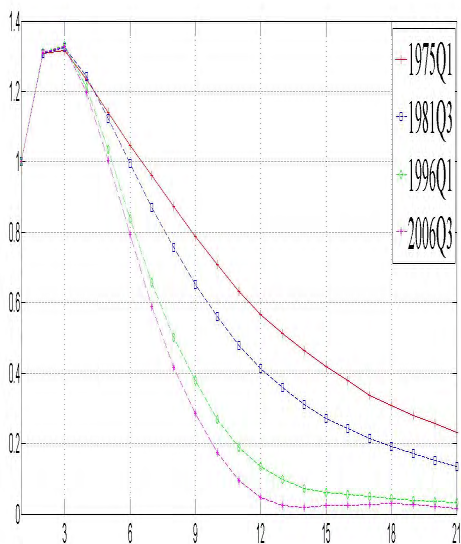


(c)

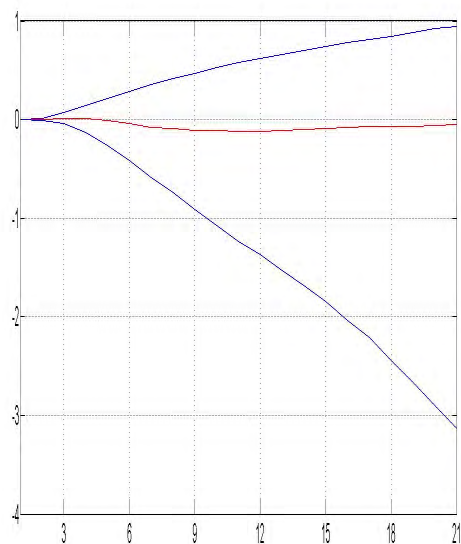


(d)

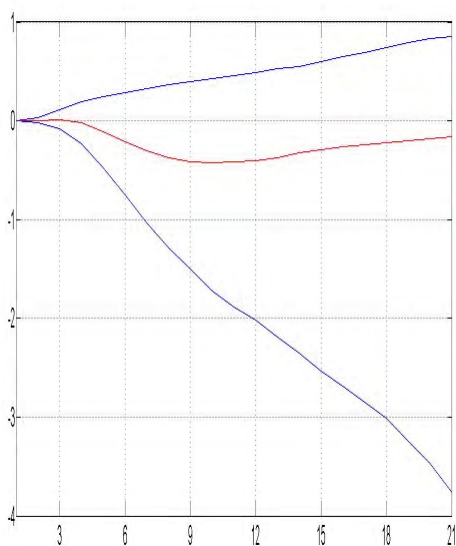
Figure 3: Response of the unemployment rate to a 1 percent monetary policy shock for 1975Q1, 1981Q3, 1996Q1 and 2006Q3: (a) medians of impulse responses; (b) response in 1981Q3 minus response in 1975Q1 with 90% credible interval; (c) response in 1996Q1 minus response in 1975Q1 with 90% credible interval; (d) response in 2006Q3 minus response in 1996Q1 with 90% credible interval.



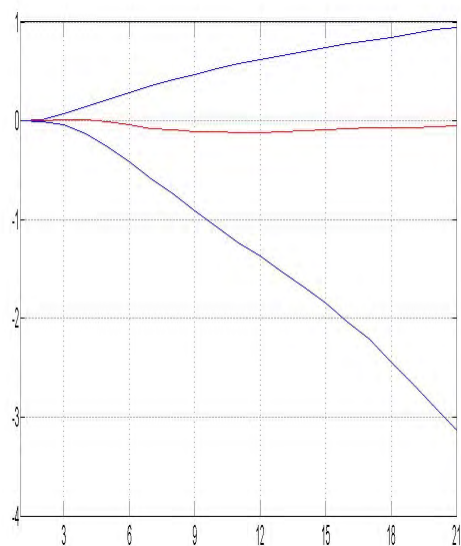
(a)



(b)



(c)



(d)

Figure 4: Response of the interest rate to 1 percent monetary policy shock for 1975Q1, 1981Q3, 1996Q1 and 2006Q3: (a) medians of impulse responses; (b) response in 1981Q3 minus response in 1975Q1 with 90% credible interval; (c) response in 1996Q1 minus response in 1975Q1 with 90% credible interval; (d) response in 2006Q3 minus response in 1996Q1 with 90% credible interval.

support from the data, implying that it is the greater flexibility in how the VAR parameters change that makes a difference in shaping impulse responses. Specifically, because the BEQ model is preferred to TVP and SB models, we argue that the impulse responses derived from this model provide better estimates of the effects of a monetary policy shock. These estimates suggest a weaker response of inflation to a monetary policy shock since the 1980s.

7 The Natural Rate of Unemployment and the Short Run Phillips Curve

7.1 Dynamics of the Natural Rate of Unemployment

Following Milton Friedman’s (1968) presidential address to the American Economic Association, the natural rate of unemployment (NRU) and the related concept of the non-accelerating inflation rate of unemployment (NAIRU) have been central concepts in macroeconomic modeling. Traditional approaches to estimating the natural rate often impose some restrictions to make the natural rate constant or, at most, allowing a few discrete jumps at certain periods of time (see Papell et al., 2000), force the NRU to be a function of time using a “spline” (see Staiger et al., 1997), or other techniques such as calibrated unobserved-components models (see Gordon, 1997), low-pass filtering (see Staiger et al., 2001) and the Hodrick-Prescott filter (see Ball and Mankiw, 2002). King and Morley (2007) endogenize the NRU as the steady-state derived from a structural vector autoregression (SVAR) in the spirit of the following quote from Phelps (1994):

“In a useful shorthand one may characterize the theory here as endogenizing the natural unemployment rate – defined now as the current equilibrium steady-state rate, given the current capital stock and any other state variables.”

Hence, the natural rate of unemployment is not necessarily a constant. Instead, King and Morley (2007) estimate a time-varying steady-state of unemployment following Beveridge and Nelson (1981) by calculating the long-run forecast in levels $\bar{y}_t = \lim_{h \rightarrow \infty} E_t y_{t+h}$ conditional on the information set available at time t . We follow this strategy by first casting the VAR into its companion form:

$$Y_{t+1} = g_t + F_t Y_t + \varepsilon_{Y,t+1}.$$

Then, we use the companion form to calculate forecasts by assuming that VAR parameters remain constants at their current values as time goes forward—i.e., in each period of time, a time-invariant VAR is assumed based on the time-varying parameter estimates for that

period.¹⁴ Conditional on the information set available at time t , g_t, F_t , the long-run forecast is

$$\bar{Y}_t = \lim_{h \rightarrow \infty} Y_{t+h} = \lim_{h \rightarrow \infty} \left(\sum_{k=0}^h F_t^k g_t + F_t^h Y_t \right) \quad \text{and} \quad \bar{u}_t = s_u \bar{Y}_t, \quad (6)$$

where s_u is a selector vector for the unemployment rate and \bar{u}_t is the time-varying NRU.¹⁵

The point estimates (based on posterior medians in each period) of the natural rate of unemployment, the 68% credible intervals, and the actual unemployment rate are plotted in Figure 5. The point estimates for the NRU range from 3.7 – 7.8 percent, which is less volatile compared to the range of 1.8 – 9.5 percent of the point estimate obtained by King and Morley (2007), but is comparable to Phelps’s (1994) estimates. The uncertainty of the point estimate has declined since the mid-1980s, which may be due to the substantial decline in the volatility of exogenous shocks around that time. Besides the difference in the range of the NRU, our estimate is smoother than that derived in King and Morley (2007). This is possibly because our model allows any block of the VAR coefficients to stay constant at their previous values, so that the trend unemployment is not forced to drift as a random walk in each period of time, and the magnitude of drift can be small if only some of VAR coefficients change.

7.2 Test of the Short-Run Phillips Curve

Because we have estimated the natural rate of unemployment, we can construct cyclical unemployment and test for the existence of the short-run Phillips curve. Following Gordon (1997), we investigate a “triangle” model (although without explicit supply shocks). Specifically, inflation is regressed on four lags of inflation and current cyclical unemployment u_t^C as follows:¹⁶

$$\pi_t = \sum_{i=1}^4 \delta_i \pi_{t-i} + \beta u_t^C + \omega_t, \quad E(\omega_t \pi_{t-i}) = 0, \quad E(\omega_t u_t^C) = 0, \quad \forall t. \quad (7)$$

Table 5 reports regression results based on the full sample and two subsamples.¹⁷ Two

¹⁴This assumption is common in the literature on bounded rationality and learning (see the “anticipated-utility” model in Kreps, 1998).

¹⁵Though we do not impose stationarity on F_t , it turns out that a large fraction of draws satisfy the stationary conditions, making the NRU well behaved. Note that the NRU is allowed to vary with time in both King and Morley (2007) and our paper, but they fit the whole sample by a time-invariant VAR model, albeit allowing for a unit root in the unemployment rate. By contrast, we allow the VAR coefficients and variances of error terms to be time-varying.

¹⁶The cyclical unemployment is constructed by subtracting the median NRU from the actual unemployment rate.

¹⁷Results for this regression are robust to also including an intercept or lags of cyclical unemployment.

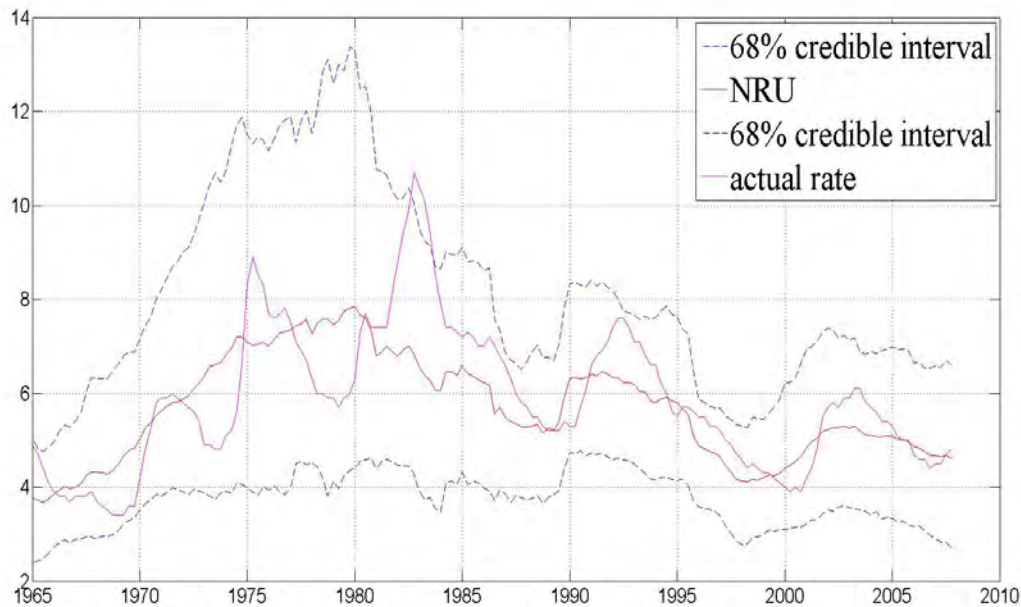


Figure 5: Natural rate of unemployment: posterior median and 68% credible interval

things stand out: First, along the lines of a Solow-Tobin test (see Solow, 1968, and Tobin, 1968), we might consider the natural rate hypothesis (i.e., a vertical long-run Phillips curve) by testing whether the sum of δ 's is significantly less than 1. Of course, as famously pointed out by Sargent (1971), the Solow-Tobin test is only informative about the natural rate hypothesis when inflation contains a unit root. As reported in Table 4, the 95% confidence intervals for the sum of the δ 's always contain 1. Hence, the natural rate hypothesis is supported by the data.¹⁸ Second, there is strong evidence supporting the existence of the short-run trade-off between inflation and the cyclical unemployment, with the short-run trade-off significant at the 5% level and estimated at $\beta = -0.4446$ for the full sample. However, there is evidence that this short-run trade-off has weakened (the median of β shifts from -0.4843 in the first subsample to -0.0428 in the latter one) and possibly even disappeared, with the 95% confidence interval for β of $[-0.3827, 0.2972]$ for the latter subsample, in accordance with the findings of Atkeson and Ohanian (2001).

In addition to the “triangle” regression analysis, we investigate time variation in the

Meanwhile, the timing of the subsamples is chosen based on the discussion in Atkeson and Ohanian (2001) and King and Morley (2007), amongst others, about a possible change in the slope of the Phillips curve in 1990.

¹⁸The decrease in the estimated sum of the δ 's in the latter subsample is likely due to the decline in the persistence of inflation, as discussed in Cogley and Sargent (2001).

Table 5: Phillips curve regression results: OLS estimates and 95% confidence intervals

Parameters	Full Sample		1965:Q1 - 1990:Q4		1991:Q1 - 2007:Q4	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
δ_1	0.5629	[0.4139, 0.7119]	0.5925	[0.3999, 0.7851]	0.3542	[0.1054, 0.6030]
δ_2	0.0811	[-0.0907, 0.2528]	0.0572	[-0.1685, 0.2829]	0.1809	[-0.0737, 0.4354]
δ_3	0.1471	[-0.0246, 0.3188]	0.1101	[-0.1158, 0.3363]	0.3041	[0.0613, 0.5469]
δ_4	0.1984	[0.0477, 0.3492]	0.2281	[0.0367, 0.4196]	0.1120	[-0.1520, 0.3760]
β	-0.4446	[-0.6157, -0.2734]	-0.4843	[-0.7007, -0.2679]	-0.0428	[-0.3827, 0.2972]
δ^*	0.9895	[0.6672, 1.3119]	0.9881	[0.5689, 1.4074]	0.9512	[0.4459, 1.4565]

The sum of the δ 's is given by $\delta^* = \sum \delta_j$, $j = 1, \dots, 4$. CI denotes "confidence interval".

short-run trade-off between inflation and cyclical unemployment using the impulse responses discussed in the previous section. Specifically, we consider the ratio of the 0 - 4 quarter average response of inflation relative to the 0 - 4 quarter average response of the unemployment rate for each structural shock.

Figure 6 plots the posterior medians of the ratios of the inflation and unemployment rate responses for each structural shock. The short-run trade-offs vary across the structural shocks and across time. The posteriors are generally quite wide and include zero, except for a shock to the unemployment rate, for which the ratio is always negative and significant up until the mid-1990s based on 68% credible intervals, as reported in Figure 7. This trade-off strengthened until around 1977 and then weakened and possibly disappeared by the mid-1990s, consistent with our findings based on the OLS regressions reported in Table 5.

Figure 8 presents a related measure of the decline in the short-run trade-off since the 1990s. This figure plots the 95% joint credible set for the ratios of the inflation and unemployment rate responses to a structural shock to the unemployment rate based on estimates from time periods A and B and conditional on a negative simulated ratio in period A. The periods for comparison that we consider in the four panels are 1975:Q1 vs. 1990:Q3, 1975:Q1 vs. 1991:Q1, 1975:Q1 vs. 2000:Q1 and 1990:Q3 vs. 2000:Q1. These are based on key business cycle reference dates of trough, peak, trough, and normal time for the four respective dates. The results evident in Figure 8 can also be summarized by the statistic F_A^B , which is defined as the fraction of simulated ratios that are greater in period B than in period A. For example, consider $F_{1975}^{1990} = 73.17\%$. This means that 73.17% of the simulated ratios in 1990:Q3 are greater than those in 1975:Q1. The equivalent statistics for the other dates are $F_{1975}^{1991} = 85.47\%$, $F_{1975}^{2000} = 91.35\%$, and $F_{1990}^{2000} = 81.41\%$. Thus, from Figure 8 and the corresponding F_A^B statistic, we can conclude that the short-run trade-off between inflation and cyclical unemployment has declined significantly since the beginning of the 1990s.

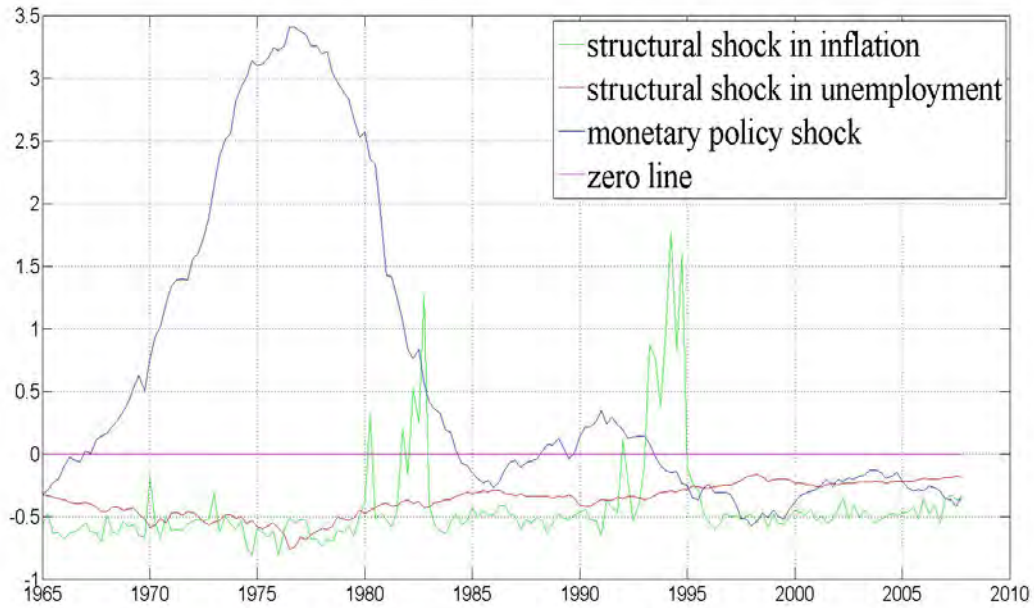


Figure 6: The posterior medians of the ratios of the inflation and unemployment rate responses (averaged over the 0 - 4 quarter horizon) for each structural shock

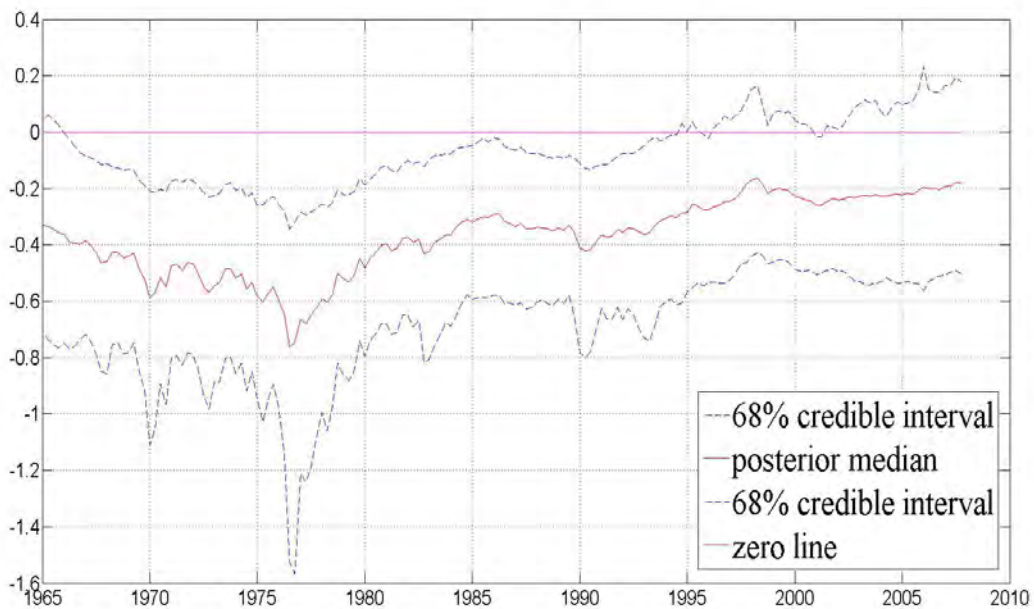


Figure 7: The posterior median and 68% credible interval for the ratio of the inflation and unemployment rate responses (averaged over the 0 - 4 quarter horizon) for a structural shock to the unemployment rate

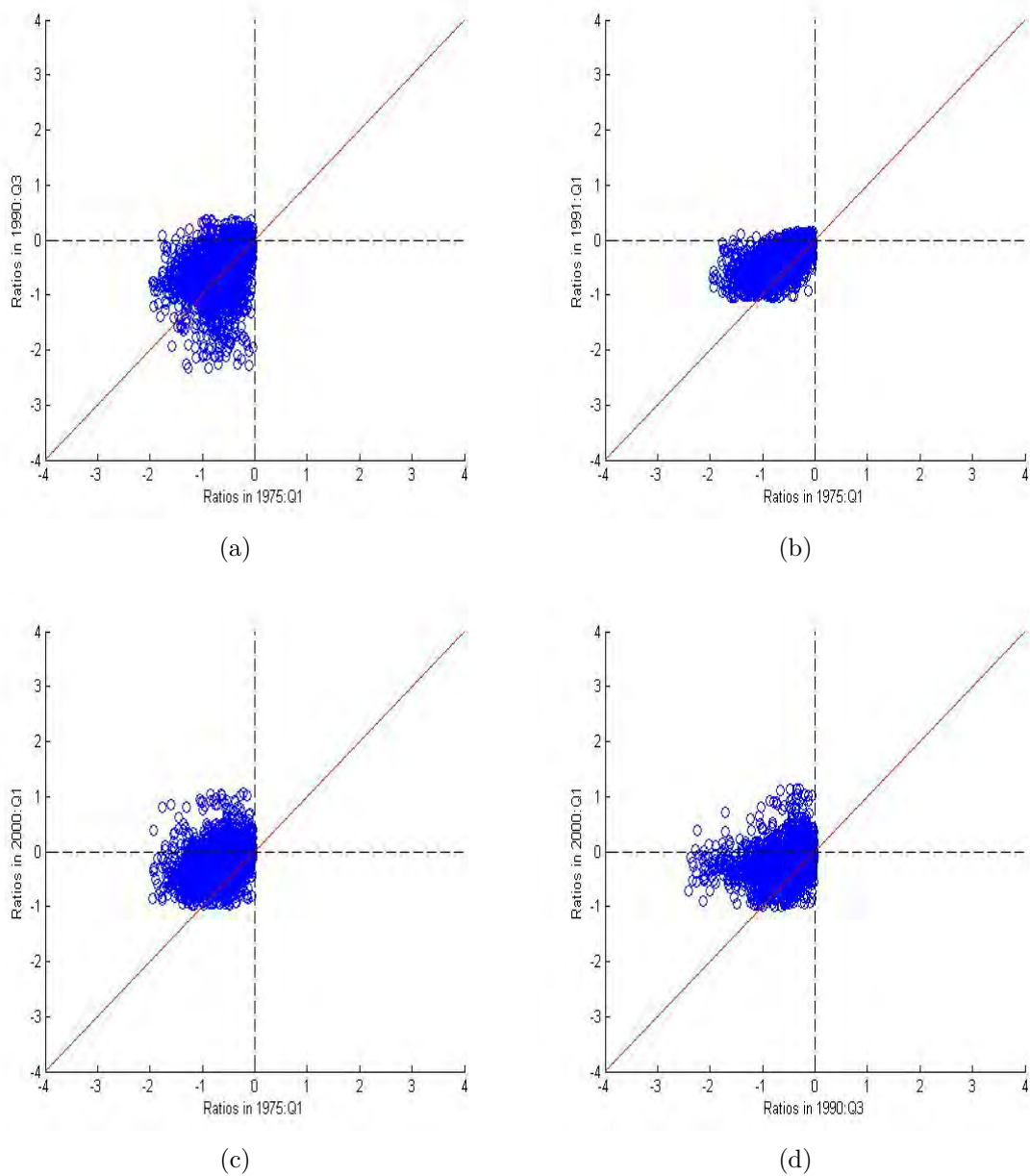


Figure 8: 95% joint credible sets of ratios of inflation and unemployment rate responses (averaged over the 0 - 4 quarter horizon) for a structural shock to the unemployment rate across certain periods of time: (a) 1975Q1 vs. 1990Q3; (b) 1975Q1 vs. 1991Q1; (c) 1975Q1 vs. 2000Q1; (d) 1990Q3 vs. 2000Q1. 95% joint credible sets are constructed by excluding 2.5% equal-tailed draws from the two marginal distributions.

8 Conclusion

In this paper, we have developed a stochastic volatility time-varying parameter vector autoregressive model with mixture innovations parameters and allowing different blocks of parameters to change at different points of time. We find that this model fits the U.S. macroeconomic data better than models that assume continuous or infrequent change in all of the model parameters at the same time. As a by-product of the flexible variation allowed in the VAR parameters, we do not force non-policy parameters to change at the same time as those related to monetary policy. This allows us to test the empirical relevance of Lucas critique. Our test provides evidence against the Lucas critique in terms of shifts in both the VAR parameters and error terms. However, the instability in the reduced-form VAR parameters is highly structured such that the parameters in the reduced-form equation for the unemployment rate vary infrequently over time, while other blocks of parameters change much more frequently.

We study the evolution of the impulse responses of inflation and the unemployment rate to monetary policy shocks as well. There is mild evidence supporting diminished effects of monetary policy after the 1970s, with an apparent shift in the response of inflation to monetary policy shock at the 3 - 9 quarter horizon only. The natural rate of unemployment is estimated as the long-run forecast of the “local-to-date” steady-state unemployment rate. Based on this measure, there is strong evidence of the existence of a short-run trade-off between inflation and cyclical unemployment. However, this trade-off has weakened since the late 1970s and has even disappeared since early 1990s.

References

- Atkeson, A. and L. Ohanian (2001), “Are Phillips Curves Useful for Forecasting Inflation?”, *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 25, 2 - 11.
- Ball, L. and N. G. Mankiw (2002), “The NAIRU in Theory and Practice”, *Journal of Economic Perspectives*, Vol. 16, 115 - 136.
- Beveridge, S. and C. R. Nelson (1981), “A 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*, Vol. 7, 151 - 174.
- Boivin, J. and M. Giannoni (2006), “Has Monetary Policy Become More Effective?”, *Review of Economics and Statistics*, Vol. 88, 445 - 462.
- Carlin, B. and T. Louis (2000), “Bayes and Empirical Bayes Methods for Data Analysis”, Boca Raton, Fla: Chapman and Hall/CRC Press.
- Carter, C. and R. Kohn (1994), “On Gibbs Sampling for State Space Models”, *Biometrika*, Vol. 81, 541 - 553.
- Cogley, T., G. Primiceri and T. Sargent (2010), “Inflation-Gap Persistence in The US”, *American Economic Journal: Macroeconomics*, Vol. 2, 43 - 69.
- Cogley, T. and T. Sargent (2001), “Evolving Post-World War II US Inflation Dynamics”, In *NBER Macroeconomics Annual*, Vol. 16, Ben S. Bernanke and Kenneth Rogoff (ed.), 331 - 388. Cambridge, MA: MIT Press.
- (2005), “Drifts and Volatilities: Monetary Policies and Outcomes in The Post WWII US”, *Review of Economic Dynamics*, Vol. 8, 262 - 302.
- Collard, F. and F. Langot (2001), “Structural Inference and The Lucas Critique”, *Annales d'économie et de statistique*, Vol. 67/68, 183 - 206.
- Cook, T. (1989), “Determinants of The Federal Funds Rate: 1979-1982”, *Federal Reserve Bank of Richmond Economic Review*, Vol. 75, 3 - 19.
- Estrella, A. and J. Fuhrer (2003), “Monetary Policy Shifts and The Stability of Monetary Policy Models”, *Review of Economics and Statistics*, Vol. 85, 94 - 104.
- Fevero, C. and D. Hendry (1992), “Testing The Lucas Critique: A Review”, *Econometric Reviews*, Vol. 11, 265 - 306.

- Friedman, M. (1968), "The Role of Monetary Policy", *American Economic Review*, Vol. 58, 1 - 17.
- Gerlach, R., C. Carter and R. Kohn (2000), "Efficient Bayesian Inference for Dynamic Mixture Models", *Journal of the American Statistical Association*, Vol. 95, 819 - 828.
- Gordon, R. (1997), "The Time-varying NAIRU and Its Implications for Economic Policy", *Journal of Economic Perspectives*, Vol. 11, 11 - 32.
- Inoue, A. and B. Rossi (2011), "Identifying The Sources of Instabilities in Macroeconomic Fluctuations", *Review of Economics and statistics*, Vol. 93, 1186 - 1204.
- Kim, S., N. Shephard and S. Chib (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models", *Review of Economic Studies*, Vol. 65, 361 - 393.
- King, T. and J. Morley (2007), "In Search of The Natural Rate of Unemployment", *Journal of Monetary Economics*, Vol. 54, 550 - 564.
- Kreps, D. (1998), "Anticipated Utility and Dynamic Choice.", In *Frontiers of Research in Economic Theory*, Donald P. Jacobs, Ehud Kalai, Morton I. Kamien and Nancy L. Schwartz (ed.). Cambridge, UK: Cambridge University Press.
- Koop, G., R. Leon-Gonzalez and R. Strachan (2009), "On The Evolution of The Monetary Policy Transmission Mechanism", *Journal of Economic Dynamics and Control*, Vol. 33, 997 - 1017.
- Kuttner, K. and P. Mosser (2002), "The Monetary Transmission Mechanism in The United States: Some Answers and Further Questions", *Federal Reserve Bank of New York, Economic Policy Review, May*, 15 - 26.
- Linde, J. (2001), "Testing for The Lucas Critique: A Quantitative Investigation", *American Economic Review*, Vol. 91, 986 - 1005.
- Lubik, T. and P. Surico (2010), "The Lucas Critique and The Stability of Empirical Models", *Journal of Applied Econometrics*, Vol. 25, 177 - 194.
- Lucas, R. (1976), "Econometric Policy Evaluation: A Critique", In *The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy*, K. Brunner and A. Meltzer (ed.), Vol. 1, 19 - 46.
- Lucas, R. and T. Sargent (1981), "After Keynesian Macroeconomics", in *Rational Expectations and Econometric Practice*, University of Minnesota Press, 295-319.

- Papell, D., C. Murray and H. Ghiblawi (2000), "The Structure of Unemployment", *Review of Economics and Statistics*, Vol. 82, 309 - 315.
- Phelps, E. (1994), "Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest, and Assets", Cambridge, MA: Harvard University Press.
- Primiceri, G. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy", *Review of Economic Studies*, Vol. 72, 821 - 852.
- Rotemberg, J. and M. Woodford (1997), "An Optimization-Based Econometric Framework for The Evaluation of Monetary Policy", In *NBER Macroeconomics Annual*, Vol. 12, Ben S. Bernanke and Julio Rotemberg (ed.), 297 - 361. Cambridge, MA: MIT Press.
- Rudebusch, G. (2005), "Assessing The Lucas Critique in Monetary Policy Models", *Journal of Money, Credit and Banking*, Vol. 37, 245 - 272.
- Sargent, T. (1971), "A Note on The Accelerationist Controversy", *Journal of Money, Credit and Banking*, Vol. 8, 721 - 725.
- (1999), "The Conquest of American Inflation", Princeton, NJ: Princeton University Press.
- Sims, C. (1992), "Interpreting The Macroeconomic Time Series Facts: The Effects of Monetary Policy", *European Economic Review*, Vol. 36, 975 - 1000.
- Sims, C. and T. Zha (2006), "Were There Regime Switches in US Monetary Policy?", *American Economic Review*, Vol. 96, 54 - 81.
- Staiger, D., J., Stock and M. Watson (1997), "How Precise Are Estimates of The Natural Rate of Unemployment?", In *Reducing Inflation: Motivation and Strategy*, Christina Romer and David Romer (ed.), 195 - 242. Chicago: University of Chicago Press.
- (2001), "Prices, wages and the US NAIRU in the 1990s", In *The Roaring Nineties: Can Full Employment Be Sustained?*, Alan Krueger and Robert Solow (ed.), 3 - 60. New York: The Russell Sage Foundation and The Century Foundation Press.
- Solow, R. (1968), "Recent Controversy on The Theory of Inflation: An Eclectic View", In *Proceedings of a Symposium on Inflation: Its Causes, Consequences, and Control*, S. Rousseaus (ed.). New York: New York University.
- Tobin, J. (1968) Discussion. In *Proceedings of a Symposium on Inflation: Its Causes, Consequences, and Control*, S. Rousseaus (ed.). New York: New York University.

Technical Appendix: The Markov Chain Monte Carlo (MCMC) Algorithm for Simulating the Posterior Density

Appendix A. Simulating $p(\theta^T, \alpha^T, \sigma^T, Q, S, W, K^T, k_5^T, k_6^T, \lambda|y^T)$

In order to simulate the joint posterior density $p(\theta^T, \alpha^T, \sigma^T, Q, S, W, K^T, k_5^T, k_6^T, \lambda|y^T)$, where T is the sample size, $\lambda = \{\lambda_{1j}, \lambda_{2j}\}_{j=1}^6$ and $K^T = (k_1^T, k_2^T, k_3^T, k_4^T)$, we draw from full conditionals, except for drawing K^T, k_5^T, k_6^T which are based on reduced conditional sampling algorithm suggested by Gerlach et al. (2000), step by step as follows.

1. Drawing latent variables K^T, k_5^T and k_6^T

In the first step, latent variables $K^T = (k_1^T, k_2^T, k_3^T, k_4^T)$ are drawn from the Gaussian linear state-space model (1) and (2). Note that

$$y_t = X_t' \theta_t + A_t^{-1} \varepsilon_t, \quad (\text{A.1})$$

$$\theta_t = \theta_{t-1} + K_t \xi_t. \quad (\text{A.2})$$

Remember that K_t has two specifications with respect to restrictions on the reduced VAR parameters according to equations or variables, but it suffices to present the simulation procedures with only one unified notation. To draw K_t , we resort to the reduced conditional sampling algorithm developed by Gerlach et al. (2000) which integrates the states out and draws K_t without conditioning on the states. This algorithm greatly improves efficiency especially when the states θ_t and K_t are highly correlated as usually the case. K_t can be drawn from

$$\begin{aligned} & p(K_t|y^T, \alpha^T, \sigma^T, Q, S, W, K_{\setminus t}, k_5^T, k_6^T, \lambda) \\ &= p(K_t|y^T, \alpha^T, \sigma^T, Q, \lambda) \\ &\propto p(y^T|K^T, \alpha^T, \sigma^T, Q, \lambda) p(K_t|\alpha^T, \sigma^T, Q, \lambda) \\ &\propto p(y^{t+1,T}|y^{1,t}, K^T, \alpha^T, \sigma^T, Q) p(y_t|y^{1,t-1}, K^{1,t}, \alpha^T, \sigma^T, Q) p(K_t|\lambda) \\ &\propto p(y^{t+1,T}|y^{1,t}, K^T, \alpha^T, \sigma^T, Q) p(y_t|y^{1,t-1}, K^{1,t}, \alpha^T, \sigma^T, Q) \times \prod_{j=1}^4 p(k_{jt}|\lambda), \end{aligned}$$

where $K_{\setminus t} = K^T \setminus K_t$ and $x^{s,t} = (x_s, x_{s+1}, \dots, x_t)$, $s < t$. The terms $p(k_{jt}|\lambda)$, $j = 1, 2, 3, 4$, can be easily obtained from the hierarchical priors. To evaluate $p(y^{t+1,T}|Y^{1,t}, K^T, \alpha^T, \sigma^T, Q)$ and $p(y_t|y^{1,t-1}, K^{1,t}, \alpha^T, \sigma^T, Q)$, please see the details in Gerlach et al. (2000).

As for k_{5t} and k_{6t} , we adapt the algorithm of Gerlach et al. (2000) to another two state-space models and draw k_{5t} and k_{6t} separately. Specifically, under the assumption that S is block diagonal, k_{5t} is drawn from the Gaussian linear state-space model with respect to the state α_t derived in Primiceri (2005):

$$\hat{y}_t = D_t \alpha_t + \Sigma_t \epsilon_t, \quad (\text{A.3})$$

$$\alpha_t = \alpha_{t-1} + k_{5t} \eta_t, \quad (\text{A.4})$$

where $D_t = \begin{bmatrix} 0 & 0 & 0 \\ -\hat{y}_{1,t} & 0 & 0 \\ 0 & -\hat{y}_{1,t} & -\hat{y}_{2,t} \end{bmatrix}$. On the other hand, following Kim et al. (1998), consider the non-Gaussian linear state-space model with respect to the state $h_t = \ln \sigma_t$ as follows:

$$y_t^{**} = 2h_t + e_t, \quad (\text{A.5})$$

$$h_t = h_{t-1} + k_{6t} \zeta_t, \quad (\text{A.6})$$

where $y_t^{**} = [y_{1t}^{**} \ y_{2t}^{**} \ y_{3t}^{**}]'$, $y_t^{**} = \ln[(y_t^*)^2 + c]$, $y_t^* = A_t(y_t - X_t' \theta_t)$, $c = 0.001$ and $e_t = [e_{1t} \ e_{2t} \ e_{3t}]'$ in which e_{jt} , $j = 1, 2, 3$ are log-chi-square distributed. Based on a mixture of normals approximation of e_{jt} 's log-chi-square distribution, we can approximate A.5 and A.6 to a sound precision by a Gaussian linear state-space model from which k_{6t} can be drawn by adapting reduced conditional sampling algorithm of Gerlach et al. (2000). Please see the details of the mixture of normals approximation in the section of "Drawing stochastic volatility σ^T ".

2. Drawing parameters of Beta priors λ

Denote the Beta priors for $p_j, j = 1, 2, \dots, 6$ by $Beta(\underline{\lambda}_{1j}, \underline{\lambda}_{2j})$, therefore, the posterior distribution of p_j is $Beta(\bar{\lambda}_{1j}, \bar{\lambda}_{2j})$, where ¹⁹

$$\bar{\lambda}_{1j} = \underline{\lambda}_{1j} + \sum_{t=1}^T k_{jt} \quad \text{and} \quad \bar{\lambda}_{2j} = \underline{\lambda}_{2j} + T - \sum_{t=1}^T k_{jt}.$$

3. Drawing reduced VAR parameters θ^T

Conditional on $y^T, \alpha^T, \sigma^T, Q, S, W, K^T, k_5^T, k_6^T, \lambda$, the states θ^T can be drawn from the state-space model A.1 and A.2 by Gibbs sampling developed in Carter and Kohn (1994). Note

¹⁹Hereafter, the underlined parameters stand for the parameters of priors and the overlined parameters represent the parameters of posteriors.

that

$$\begin{aligned}
& p(\theta^T | \alpha^T, \sigma^T, Q, S, W, K^T, k_5^T, k_6^T, \lambda, y^T) \\
&= p(\theta^T | \alpha^T, \sigma^T, Q, y^T, K^T) \\
&= p(\theta_T | \alpha^T, \sigma^T, Q, y_T, K_T) \prod_{t=1}^{T-1} p(\theta_t | \theta_{t+1}, y^t, \alpha^T, \sigma^T, Q, K_{t+1}),
\end{aligned}$$

where

$$\begin{aligned}
\theta_t | \theta_{t+1}, y^t, \alpha^T, \sigma^T, Q, K_{t+1} &\sim N(\theta_{t|t+1}, P_{t|t+1}), \\
\theta_{t|t+1} &= E(\theta_t | \theta_{t+1}, y^t, \alpha^T, \sigma^T, Q, K_{t+1}), \\
P_{t|t+1} &= Var(\theta_t | \theta_{t+1}, y^t, \alpha^T, \sigma^T, Q, K_{t+1}).
\end{aligned}$$

The last recursion of forward Kalman filter gives $\theta_{T|T}$ and $P_{T|T}$ from which θ_T can be simulated. Then $\theta_{t|t+1}$ and $P_{t|t+1}$, $t = 1, 2, \dots, T-1$, are obtained by backward recursions from $\theta_{T|T}$ and $P_{T|T}$. From $N(\theta_{t|t+1}, P_{t|t+1})$, we are able to simulate the smoothed estimates of θ_t , $t = 1, 2, \dots, T-1$. Please see the details of Gibbs sampling in Appendix B.

4. Drawing hyperparameter Q

Since we assume the prior of Q is the inverse-Wishart distribution $\mathbb{IW}(Q, \underline{\nu}_Q)$, hence Q^{-1} is governed by Wishart distribution as:

$$Q^{-1} \sim \mathbb{W}(\underline{Q}^{-1}, \underline{\nu}_Q).$$

Hence, the posterior for Q^{-1} conditional on other blocks is Wishart as well:

$$Q^{-1} | y^T, \theta^T, \alpha^T, \sigma^T, S, W, K^T, k_5^T, k_6^T, \lambda \sim \mathbb{W}(\overline{Q}^{-1}, \overline{\nu}_Q),$$

where

$$\overline{Q}^{-1} = \left[\underline{Q}^{-1} + \sum_{t=1}^T (\theta_{t+1} - \theta_t)(\theta_{t+1} - \theta_t)' \right]^{-1} \quad \text{and} \quad \overline{\nu}_Q = \underline{\nu}_Q + T.$$

5. Drawing covariances α^T

Reconsider the Gaussian linear state-space model A.3 and A.4 under the assumption of block-diagonal S . Since $\hat{y}_{1,t}$ is determined by exogenous identity shock ϵ_{1t} and $\sigma_{11,t}$, thus conditional on other blocks, $\hat{y}_{1,t}$ is predetermined in $\hat{y}_{2,t}$'s equation. So are $\hat{y}_{1,t}$ and $\hat{y}_{2,t}$

predetermined in $\hat{y}_{3,t}$'s equation. Therefore, α_t can be obtained by applying Kalman filter and the backward recursion equation by equation. Let $\alpha_t = [\alpha_{1,t}, \alpha_{2,t}]'$, where $\alpha_{1,t} = \alpha_{21,t}$ and $\alpha_{2,t} = [\alpha_{31,t}, \alpha_{32,t}]'$ are corresponding to different blocks in S , then the smoothed estimate of α_t is derived from

$$\begin{aligned}\alpha_{i,t} | \alpha_{i,t+1}, y^t, \theta^T, S_i, \sigma^T, k_{5,t+1} &\sim N(\alpha_{i,t|t+1}, \Lambda_{i,t|t+1}), \\ \alpha_{i,t|t+1} &= E(\alpha_{i,t} | \alpha_{i,t+1}, y^t, \theta^T, S_i, \sigma^T, k_{5,t+1}), \\ \Lambda_{i,t|t+1} &= Var(\alpha_{i,t} | \alpha_{i,t+1}, y^t, \theta^T, S_i, \sigma^T, k_{5,t+1}), \quad i = 1, 2.\end{aligned}$$

6. Drawing hyperparameter S

Recall that we separate S into two blocks S_1 and S_2 each governed by inverse-Wishart distribution $\mathbb{IW}(\underline{S}_j, \underline{\nu}_{S_j}), j = 1, 2$. Equivalently, $S_j^{-1} \sim \mathbb{W}(\underline{S}_j^{-1}, \underline{\nu}_{S_j}), j = 1, 2$. Thus, the conditional posterior for $S_j, j = 1, 2$, are as follows:

$$S_j^{-1} | y^T, \theta^T, \alpha^T, \sigma^T, Q, W, K^T, k_5^T, k_6^T, \lambda \sim \mathbb{W}(\bar{S}_j^{-1}, \bar{\nu}_{S_j}),$$

where

$$\bar{S}_j^{-1} = \left[\underline{S}_j^{-1} + \sum_{t=1}^T (\alpha_{j,t+1} - \alpha_{j,t})(\alpha_{j,t+1} - \alpha_{j,t})' \right]^{-1} \quad \text{and} \quad \bar{\nu}_{S_j} = \underline{\nu}_{S_j} + T, \quad i = 1, 2.$$

7. Drawing stochastic volatility σ^T

The stochastic volatility σ^T are drawn from the non-Gaussian linear state-space model A.5 and A.6 based on a mixture of seven normals approximation *a la* Kim et al. (1998) with component probabilities q_l , means $m_l - 1.2704$ and variances $v_l^2, l = 1, 2, \dots, 7$. Please see the constants $\{q_l, m_l, v_l^2\}$ chosen for matching a number of moments of the $\log(\chi^2(1))$ distribution in Kim et al. (1998). Note that y_{it}^{**} and y_{jt}^{**} are independent of one another for $i \neq j$, hence, e_{it} is independent of e_{jt} as well. Thus, it allows us to employ the same mixture of normals to approximate any element in e_t .

Define the state-indicator matrix $s^T = [s_1, s_2, \dots, s_T]'$, $s_t = [s_{1t}, s_{2t}, s_{3t}]'$, $s_{jt} \in \{1, 2, \dots, 7\}$, $j = 1, 2, 3$ and $t = 1, 2, \dots, T$, showing in each period of time which member of the mixture of normals is used for each element of e_t . Conditional on other blocks, after determining the members of the mixture of normals used for approximation for e_t , the system obtained is a Gaussian linear state-space model in which h_t can be easily drawn based on the standard Kalman filtering and backward recursions as in previous steps. Specifically, smoothed

estimates of h_t can be drawn recursively from

$$\begin{aligned} h_t | h_{t+1}, y^t, \theta^T, \alpha^T, W, k_6^T, s^T &\sim N(h_{t|t+1}, H_{t|t+1}), \\ h_{t|t+1} &= E(h_t | h_{t+1}, y^t, \theta^T, \alpha^T, W, k_6^T, s^T), \\ H_{t|t+1} &= Var(h_t | h_{t+1}, y^t, \theta^T, \alpha^T, W, k_6^T, s^T). \end{aligned}$$

After obtaining a new draw of h^T , s^T can be updated as in Kim et al. (1998) for each s_{jt} independently from the discrete density

$$Pr(s_{jt} = l | y_{jt}^{**}, h_{jt}) \propto q_l f_N(y_{jt}^{**} | 2h_{jt} + m_l - 1.2704, v_l^2), \quad j = 1, 2, 3, l = 1, 2, \dots, 7,$$

where $f_N(\cdot)$ stands for the normal density. Finally, the smoothed estimate of σ_t can be recovered by transform $\sigma_t = \exp\{0.5h_t\}$.

8. Drawing hyperparameter W

Note that $W \sim \mathbb{IW}(W, \underline{v}_W)$, i.e. $W^{-1} \sim \mathbb{W}(W^{-1}, \underline{v}_W)$, where $\mathbb{W}(\cdot, \cdot)$ and $\mathbb{IW}(\cdot, \cdot)$ stand for Wishart distribution and inverse-Wishart distribution, respectively. Hence, the posterior for W^{-1} conditional on other blocks reads:

$$W^{-1} | y^T, \theta^T, \alpha^T, \sigma^T, Q, S, K^T, k_5^T, k_6^T, \lambda \sim \mathbb{W}(\bar{W}^{-1}, \bar{v}_W),$$

where

$$\bar{W}^{-1} = \left[W^{-1} + \sum_{t=1}^T (h_{t+1} - h_t)(h_{t+1} - h_t)' \right]^{-1} \quad \text{and} \quad \bar{v}_W = \underline{v}_W + T.$$

Appendix B. Gibbs Sampling for State-Space Models

Cast the Gaussian linear state-space models in this paper into the following state-space form:

$$\begin{aligned} \text{Measurement equation:} \quad y_t &= F_t \beta_t + u_t, \\ \text{State equation:} \quad \beta_t &= \beta_{t-1} + v_t, \end{aligned}$$

where

$$\begin{bmatrix} u_t \\ v_t \end{bmatrix} \sim iid. N \left(\begin{bmatrix} u_t \\ v_t \end{bmatrix}, \begin{bmatrix} R_t & 0 \\ 0 & Q \end{bmatrix} \right).$$

Define

$$\begin{aligned}\beta_{t|s} &= E(\beta_t | y^s, F^s, R^s, Q), \\ P_{t|s} &= \text{Var}(\beta_t | y^s, F^s, R^s, Q).\end{aligned}$$

Given the mean and variance of the initial state, $\beta_{0|0}$ and $P_{0|0}$, the forward Kalman filter yields:

$$\begin{aligned}\beta_{t|t-1} &= \beta_{t-1|t-1}, \\ P_{t|t-1} &= P_{t-1|t-1} + Q, \\ \kappa_t &= P_{t|t-1} F_t' (F_t P_{t|t-1} F_t' + R_t)^{-1}, \\ \beta_{t|t} &= \beta_{t|t-1} + \kappa_t (y_t - F_t \beta_{t|t-1}), \\ P_{t|t} &= P_{t|t-1} - \kappa_t F_t P_{t|t-1}.\end{aligned}$$

After obtaining $\beta_{T|T}$ and $P_{T|T}$, we draw β_T from $N(\beta_{T|T}, P_{T|T})$. Then the draw of β_T and the output derived from the above forward Kalman filter are used for backward recursion as follows:

$$\begin{aligned}\beta_{t|t+1} &= \beta_{t|t} + P_{t|t} P_{t+1|t}^{-1} (\beta_{t+1} - \beta_{t|t}), \\ P_{t|t+1} &= P_{t|t} - P_{t|t} P_{t+1|t}^{-1} P_{t|t},\end{aligned}$$

which provide $\beta_{T-1|T}$ and $P_{T-1|T}$ that are used to generate β_{T-1} . Likewise, $\beta_{T-2}, \beta_{T-3}, \dots, \beta_1$ are drawn from $N(\beta_{T-2|T-1}, P_{T-2|T-1}), N(\beta_{T-3|T-2}, P_{T-3|T-2}), \dots, N(\beta_{1|2}, P_{1|2})$, respectively.