Australian School of Business Research Paper No. 2013 ECON15
First draft June 2013
Australian School of Business Research Paper No. 2013 ECON15A
This version March 2014

Structural Evolution of the Postwar U.S. Economy

Yuelin Liu
James Morley

Structural Evolution of the Postwar U.S. Economy

Yuelin Liu*
University of New South Wales

James Morley
University of New South Wales

March 7, 2014

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 tradeoff 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 two anonymous referees, Adrian Pagan, Barbara Rossi, Tara Sinclair, Andrea Tambalotti, Peter Tulip, and participants in the 18th Australasian Macroeconomics Workshop, the Econometric Society Australasian Meeting 2013, and the 2013 RBA Quantitative Macroeconomics Workshop, especially our discussants Joshua Chan and Barry Rafferty, for helpful comments and suggestions. We also thank Todd Clark for pointing out a recent correction by Primiceri and Del Negro (2013) to Primiceri (2005). 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 at least some parameters to change over time. In addition, a substantial decline in the volatility of macroeconomic variables, often referred to as "the Great Moderation", has occurred since the mid-1980s. Together, these changes imply that a conventional vector autoregressive (VAR) model with constant parameters and homoskedastic shocks is inadequate for the postwar U.S. data.

In order to allow for changes in model parameters 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 abruptly between a few (possibly recurrent) regimes for conditional mean and/or variance parameters, where 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 and/or variance parameters (see, for example, Cogley and Sargent (2001, 2005), Primiceri (2005), and Cogley et al. (2010)).

Recently, a few models that bridge the MS and TVP approaches have been proposed. Koop and Potter (2007) develop a non-reversible change-point model with unknown number of regime shifts and Poison-distributed regime durations, while Giordani and Kohn (2008) introduce an alternative flexible framework called a 'dynamic mixture model' or 'mixture innovation model' in which the timing of regime shifts is determined by a latent variable and subgroups of model parameters are allowed to vary independently.[1] Building on these studies, Koop et al. (2009) apply the mixture innovations framework to extend the standard TVP-VAR of Primiceri (2005). They do so by introducing independent binary latent variables for three blocks of parameters corresponding to conditional mean parameters, variances, and the contemporaneous cross-equation impacts of shocks, allowing the data to determine the occurrence of a regime shift for each block in each period of time. However, principal components analysis of the variance-covariance matrix governing the magnitude of changes in TVP-VAR parameters conducted by Cogley and Sargent

[1]Notably, models with mixture innovations often have improved forecasting performance relative to simpler models, supporting their usefulness in describing the time series properties of the macroeconomic data. For example, Giordani and Villani (2010) forecast nine quarterly macroeconomic series from the United States, Sweden, and Australia using a mixture innovation model and find it outperforms related models with restrictions such as homoskedastic errors or smooth, continuous changes in parameters. Likewise, Groen et al. (2013) find very accurate real-time point and density forecasts for a multivariate model of U.S. inflation with mixture innovations.

(2005) shows that conditional mean parameters appear to vary in a highly structured way that does not correspond to uniform changes across these parameters. Specifically, a small number of principal components explain most of the time variation in conditional mean parameters, with loadings varying considerably across parameters. Cogley and Sargent (2005) speculate that this pattern could be due to cross-equation restrictions associated with private agents' optimization and foresight in the context of adaptive learning by the policymaker, as considered in Sargent (1999). Meanwhile, if some parameters vary more frequently and more strongly, while 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.

Motivated by the possibility that not all conditional mean parameters need to change together, we extend Koop et al.'s (2009) analysis to allow for variation at different points of time in subgroups of VAR parameters, including different blocks of the conditional mean parameters. Because changes in each block of parameters are controlled by a Bernoulli distributed latent variable, the posterior density of the probability parameter for the Bernoulli distribution reflects the frequency of occurrence of breaks in a given block. Then, if the true model is the stochastic volatility TVP-VAR model, as in Primiceri (2005), the data will push the probability parameter for each block to one. Otherwise, if the true model is a MS-VAR model, the probability parameters will be much smaller than one, with differences in probability parameters across blocks 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 stochastic shifts in the different blocks of parameters, which appear to be relevant in practice according to our results.

Building on Koop et al.'s (2009) modeling strategy, our paper makes three contributions: First, because we divide the VAR parameters into "policy" and "non-policy" blocks, the frequency of changes in the non-policy blocks relative to that of the 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 time series behaviour of the macroeconomic variables affected by policy. This test is different than simulation-based approaches to testing the Lucas critique often considered in the literature; see, for example, Estrella and Fuhrer (2003), Lindé (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 error terms. Notably, 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 are no statistically significant changes in impulse responses for monetary policy shocks over the postwar period, and Kuttner and Mosser (2002) and Boivin and Giannoni (2006), who find that the effects of monetary policy on the U.S. economy have weakened since 1980s. Based on our model, we find that the effects of monetary policy on inflation have only changed over time at the 3-9 quarter horizon, while the effects on the unemployment rate appear not to have changed at any horizon. 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 tradeoff between inflation and cyclical unemployment. We find evidence of a short-run tradeoff, with some support for a nonlinear relationship that is stronger for higher levels of lagged inflation. However, the tradeoff has clearly weakened since the late 1970s and has even disappeared since the mid-1990s, coinciding with the anchoring of inflation expectations at relatively low levels in recent years.

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 provides model fit and robustness analysis. Section 5 considers the empirical relevance of Lucas critique. Section 6 reports on the evolution of impulse response functions for a monetary policy shock on inflation and the unemployment rate. Section 7 examines the natural rate of unemployment and the short-run tradeoff between inflation and cyclical unemployment. Section 8 concludes.

2 Model

One of the contributions of this paper is to broaden the number of blocks of VAR parameters linked to latent variables relative to Koop et al. (2009). This allows us to consider 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. For our analysis, we consider 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:

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

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

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 our analysis, 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 to 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 $y_t = [\pi_t \quad u_t \quad i_t]'$, 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 only respond to a monetary policy shock with at least a one-period lag.[3] In practice, structural shocks are recovered via a Cholesky decomposition of the variance-covariance matrix of the reduced-form error terms as follows:

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

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

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

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

$$X_t = I_3 \otimes [1 \; y_{t-1}' \; y_{t-2}'] \,. \tag{1}$$

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:

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

[2] A 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).

[3] It is well understood that the order of variables in a recursive identification scheme will matter given correlation between the reduced-form errors. However, our results for a monetary policy shock are largely 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 there is 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. 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 $K_t^{(1)} = diag\{K_{1t}^{(1)},\ K_{2t}^{(1)},\ K_{3t}^{(1)},\ K_{4t}^{(1)},$
$K_{2t}^{(1)},\ K_{3t}^{(1)},\ K_{4t}^{(1)}\}$ (by equations) and $K_t^{(2)} = diag\{K_{1t}^{(2)},\ K_{2t}^{(2)},\ K_{2t}^{(2)}, K_{2t}^{(2)},\ K_{2t}^{(2)},\ K_{2t}^{(2)},\ K_{2t}^{(2)}\}$ (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:

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

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

where S, W are positive definite and S is block diagonal with each block 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 the following variance-covariance matrix V:

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

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 approach greatly improves the efficiency of the sampler when K_t and θ_t are highly correlated.[4] 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.[5] See the detailed MCMC algorithm in the technical appendix.

3 Data and Priors

3.1 Data

As discussed above, we consider a small three-variable VAR model 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 impact on inflation (seasonally adjusted compounded annual rate of change of Personal Consumption Expenditures, series ID: PCECTPI) and the unemployment rate (seasonally adjusted civilian unemployment rate, all workers over 16, series ID: UNRATE).[6] The series are quarterly and run from 1954Q3 to 2007Q4, where the end of the sample is chosen to avoid the recent zero-lower-bound period that would require a more complicated model to capture U.S. monetary policy than can be easily accommodated in the mixture innovations framework.

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 10^{th} 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 for parameter draws. As shown

[4] As noted by Koop et al. (2009), modeling correlations between the k_{jt}'s would bring a huge cost in terms of increasing the computation burden. Meanwhile, 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.

[5] Recently, Del Negro and Primiceri (2013) have made a correction to the MCMC procedures in Primiceri (2005). They retain most of the procedures except that sampling of stochastic volatilities is preceded by sampling of states for mixture components approximations to errors with log chi-square distributions. We follow their updated procedures.

[6] All of the data series were downloaded from FRED managed by the Federal Reserve Bank of St. Louis at http://research.stlouisfed.org/fred2/. The results are generally robust to different measurements of inflation and short-term interest rates, for example, using the GDP deflator or the 3-month Treasury bill rate. The Personal Consumption Expenditures (PCE) deflator is used because it has been, at least in recent years, the Fed's preferred measure of the cost of living. When the 3-month Treasury bill rate is considered instead of the federal funds rate, an unemployment puzzle (i.e., a contractionary monetary policy associated with a decline in the unemployment rate) appears, as in Koop et al. (2009). Therefore, we take the federal funds rate (FFR), which is directly under the 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 officially framed its policy in terms of monetary aggregates. But Cook (1989) argues that, even in that episode, the FFR serves as a satisfactory policy indicator. Hence, we believe it is appropriate to treat the FFR as the policy instrument across most of the postwar period, except for the recent zero-lower-bound period that we do not consider in our analysis.

in Figure 1, the 20^{th}-order autocorrelations for almost all of parameters draws (including for the hyperparameters) are below 0.15 and only a few are as high as $0.2 - 0.3$. Thus, the posterior draws are mixing well and the convergence check is satisfactory.

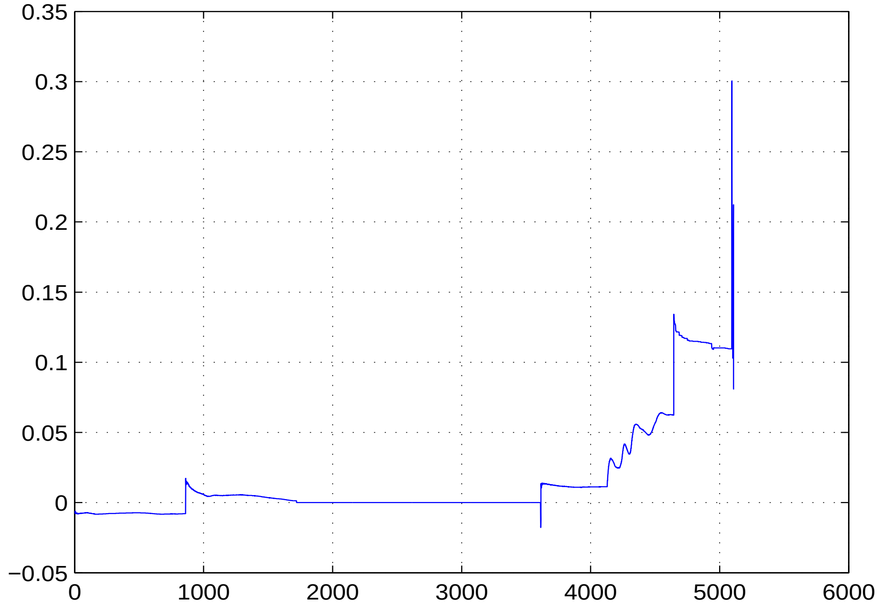


Figure 1: 20^{th}-Order Autocorrelations for Parameter Draws. From left to right, condition mean parameters θ_t from 1-3612, covariances α_t from 3613-4128, variances σ_t from 4129-4644, hyperparameters Q, S, W from 4645-5103, and probability parameters p_j from 5104-5109, where $t = 1, 2, \cdots, T$ and $j = 1, 2, \cdots, 6$.

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, 1954Q3 – 1964Q4) are employed to calibrate the priors. Specifically, a time-invariant VAR model is estimated using conditional MLE, which produces point estimates, $\hat{\theta}_0$, for the conditional mean parameters and their corresponding variances, $V(\hat{\theta}_0)$. Estimates, $\hat{\Omega}_0$, of the variance-covariance matrix for the VAR errors are obtained as well and $\hat{\alpha}_0$, $\hat{\sigma}_0$ are derived from decomposing $\hat{\Omega}_0$. The variance, $V(\hat{\alpha}_0)$, of $\hat{\alpha}_0$ is obtained by simulation from a Wishart distribution with scatter matrix $\hat{\Omega}_0$ and degree of freedom set to 40. 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 little weight as possible on prior beliefs, the degree of freedom for 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

$$Prob(k_{jt} = 1) = p_j, \qquad j = 1, 2, \cdots, 6, \tag{5}$$

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 $Beta(\lambda_{1j}, \lambda_{2j}),\ j = 1, 2, \cdots, 6$, 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 is somewhere around a 50% chance. The prior is reasonably diffuse, although it places very little weight on parameter values corresponding to SB and TVP versions of the model based on the tight priors for the p_j's suggested above. Table 1 provides a summary of our priors on $p_j, j = 1, 2, \cdots, 6$, where the label BEQ denotes the benchmark model varying with respect to equations, the label BVA denotes the benchmark model varying with respect to variables, and SBEQ, SBVA, TVPEQ and TVPVA are analogous labels for the versions of the model with strong priors on the p_j's. In general, we combine these priors for the p_j's with those for state vectors $\theta_t, \alpha_t, \sigma_t$ and hyperparameters Q, S, W provided above.

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

Table 1: Priors for $p_j = Prob(k_{jt} = 1) : Beta(\lambda_{1j}, \lambda_{2j}), j = 1, 2, \cdots, 6, \ \forall t$.

Models	λ_{1j}	λ_{2j}
BEQ/BVA	1	1
TVPEQ/TVPVA	1	0.01
SBEQ/SBVA	0.01	10

4 Model Fit and Robustness Analysis

Table 2 reports on the fit of the benchmark model, the SB model, and the TVP model using i) posterior medians of p_j's and ii) the expected value of log-likelihood function as described in Carlin and Louis (2000). Note that marginal likelihoods are difficult to calculate in this setting given the high-dimensional parameter space.

The first result that stands out in Table 2 is that the BEQ model with slopes in the same equation changing together receives strongest support based on the expected log-likelihood $E(\log L|Y)$.[8] Somewhat supportive of Primiceri (2005), Cogley and Sargent (2005), and Koop et al. (2009), the probabilities of parameter change for the BEQ model suggest that the shock volatilities have changed frequently over time with $E(p_6|Y) = 0.9790$ (standard deviation 0.0205). However, the slopes in unemployment equation appear to be relatively stable with $E(p_3|Y) = 0.0329$ (standard deviation 0.0139), which is substantially smaller than the probabilities of parameter change in other blocks. This result implies that a change in the slopes of the unemployment equation is expected to occur only once every 25 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 the VAR parameters are highly structured ,with much stronger support for blocks based on equations than based on variables.[9]

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 estimates for the benchmark model. This can be treated as robustness analysis for our modeling strategy. For example, in the SBEQ model, the

[8] The expected log-likelihood is measured 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) $x^T = (x_1, x_2, \cdots, x_T)$. Taking into account a penalty for number of parameters, the ranking of competing models is robust to the Akaike information criterion (AIC) and the Schwartz information criterion (SIC) computed at the expected likelihood. In frequentist econometrics, AIC and SIC are evaluated at the mode of likelihood function. Although the MCMC sampler does not provide a precise estimate of the posterior mode, we have checked AIC and SIC with respect to the highest likelihood values obtained from the sampler and found that the BEQ model again performs best, with the ordering of other models the same except the rankings of the BVA and Primiceri models are interchanged.

[9] As evident from Table 2, the slopes generally change less frequently for the BVA model than for the BEQ model. We note that our general findings are reasonably robust to considering blocking by variables instead of equations. However, we find less variation in impulse responses over time and the estimation uncertainty about the time-varying tradeoff between inflation and cyclical unemployment is much greater than for the BEQ model.

Table 2: 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	0.8607	0.8773	0.0329	0.8882	0.4866	0.9790	1984.2
	(0.0345)	(0.0367)	(0.0139)	(0.0364)	(0.2882)	(0.0205)	
BVA	0.7444	0.0057	0.2398	0.1221	0.4950	0.9733	1821.8
	(0.0348)	(0.0056)	(0.0324)	(0.0247)	(0.3034)	(0.0258)	
TVPEQ	0.9999	0.9999	0.0353	0.9999	0.9941	0.9998	1839.3
	(0.0008)	(0.0005)	(0.0143)	(0.0006)	(0.0492)	(0.0019)	
TVPVA	0.0058	0.9998	0.0058	0.9999	0.9935	0.9998	1269.0
	(0.0058)	(0.0015)	(0.0058)	(0.0008)	(0.0491)	(0.0017)	
SBEQ	0.3333	0.1664	0.1344	0.0956	0.0009	0.0003	896.4
	(0.0395)	(0.0289)	(0.0254)	(0.0225)	(0.0096)	(0.0029)	
SBVA	0.1144	0.1132	0.1841	0.0554	0.0010	0.0010	1137.6
	(0.0242)	(0.0235)	(0.0294)	(0.0171)	(0.0102)	(0.0075)	
KLS	0.9757				0.5047	0.9618	1846.5
	(0.0118)				(0.2892)	(0.0375)	
Primiceri	–				–	–	1784.7
	–				–	–	

Standard deviations are listed in parentheses. "KLS" represents the benchmark model in Koop et al. (2009) for which $k_{1t} = k_{2t} = k_{3t} = k_{4t}$ for all t. For "Primiceri", $k_{it} = 1$, $i = 1, 2, \cdots, 6$, for all t, as in Primiceri (2005).

parameters of $Beta(\lambda_{1j}, \lambda_{2j})$ priors on p_j's are set to $\lambda_{1j} = 0.01, \lambda_{2j} = 10, j = 1, 2, \cdots, 6$, 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 that $E(p_1|Y) = 0.3333, E(p_2|Y) = 0.1664, E(p_3|Y) = 0.1344, E(p_4|Y) = 0.0956, E(p_5|Y) = 0.0009, E(p_6|Y) = 0.0003$ with standard deviations $0.0395, 0.0289, 0.0254,$
$0.0225, 0.0096$ and 0.0029, respectively. The posterior expected probabilities with respect to the conditional mean parameters suggest that parameter changes happen approximately every 3 to 10 quarters, which strongly rejects a prior belief that very few breaks occur over time. Meanwhile, in the TVPEQ model, λ_{1j} and λ_{2j} are set to 1 and 0.01, respectively. So $E(p_j) = 0.99$ with standard deviation 0.08 clearly favors a time-varying parameter with stochastic volatility model. However, the posterior mean value of p_3 is 0.0353 with standard deviation 0.0143. Thus, the posterior probability of observing a break in the slopes in unemployment equation in every period declines substantially from the prior. This suggests that the slopes in the unemployment rate equation are more stable than the other blocks in the model and also supports the idea of highly structured changes in the VAR parameters.

As a more general robustness check, we investigate the standard deviations of the reduced-form VAR errors for three alternative models: BEQ, KLS, and Primiceri. Figure 2 plots the posterior medians of standard deviations for these errors. A decline in volatility since the mid-1980s is evident from the plots. The stochastic volatility estimates for the KLS and Primiceri models track each other closely for all three errors. Meanwhile, the stochastic volatility estimates for inflation and interest rate errors from our BEQ model are almost the same as those for the KLS and Primiceri models. By contrast, the stochastic volatility estimates for unemployment rate errors derived from our BEQ model are slightly higher than for the KLS and Primiceri models. This difference is not surprising because the conditional mean parameters in the unemployment equation vary much more frequently in the KLS and Primiceri models than in our BEQ model, as evident in Table 2. Thus, less of the overall variation in the unemployment rate is ascribed to the VAR errors for Koop et al. (2009) and Primiceri (2005) compared to our BEQ model. However, consistent with the finding by Koop et al. (2009), we find that the restrictions on the structure of time-varying features of conditional mean parameters do not fundamentally change the volatility estimates, especially in terms of general changes over time, although they do influence the impulse response functions, as will be discussed in Section 6.

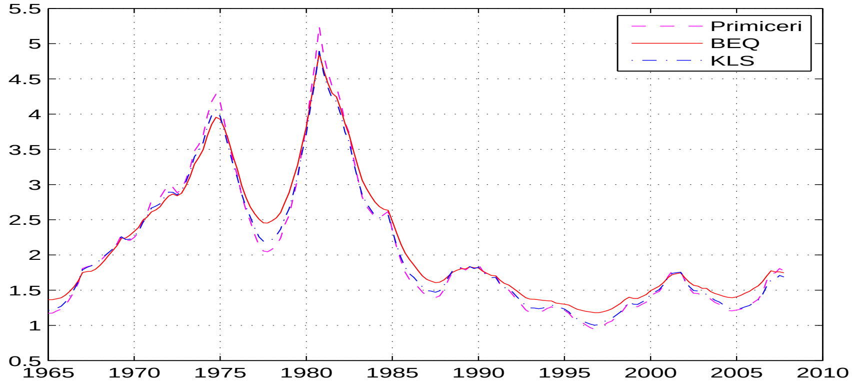
In summary, the BEQ model receives the strongest support in the model comparison, while it captures similar features in the data such as volatility levels and changes when compared with related models. Hereafter, the empirical results about the evolution of U.S. economy are based on estimates for this model.

5 Testing the Lucas Critique

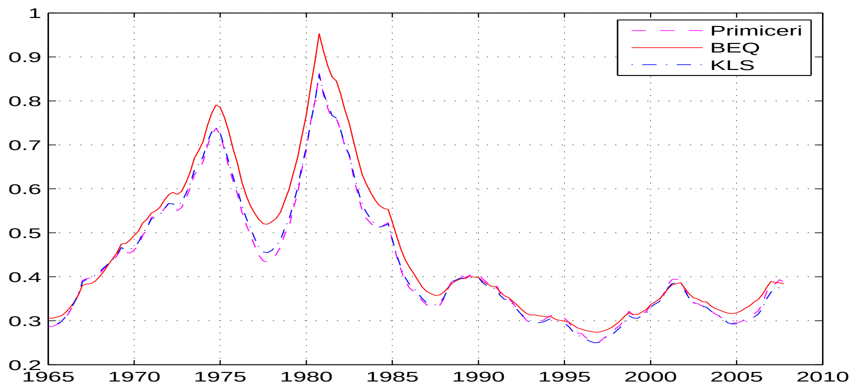
Ever since Lucas's (1976) seminal study, it has been widely-recognized by macroeconomists that reduced-form econometric models could be inappropriate for policy analysis if there are changes in parameters describing policy. However, a relatively large literature—see Fevero and Hendry (1992), Estrella and Fuhrer (2003), Lindé (2001) and Rudebusch (2005), amongst others—casts doubt on the empirical relevance of Lucas critique by considering Chow tests and superexogeneity tests. In a recent study, 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 previous studies for implicitly assuming homoskedasticity of the reduced-form error terms, undermining the power of the tests.[10]

Most of the recent studies testing the Lucas critique rely on simulating data from a specified DSGE model as if the model were the "true" data generating process (DGP) of the macro variables

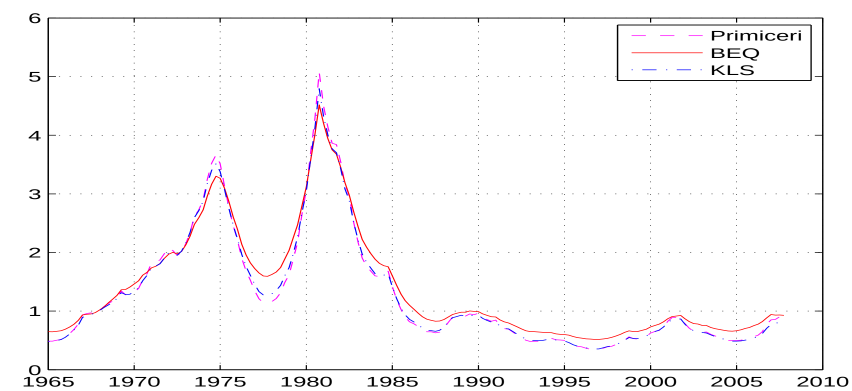
[10] Another potential issue is the low power of the superexogeneity test in small samples, as discussed in Lindé (2001) and Collard et al. (2001).



(a) Standard deviations of the error term in the inflation equation



(b) Standard deviations of the error term in the unemployment equation



(c) Standard deviations of the error term in the interest rate equation

Figure 2: Posterior medians of standard deviations of error terms in observation equations.

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" and "non-policy" 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.[11] But this is less of a problem than it may at first appear because a change in parameters for the structural policy equation should induce a change in the parameters for the reduced-form policy equation. Then, if the Lucas critique holds, the change in the policy parameters should also induce a change 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 change 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 change in the reduced-form policy parameters without a corresponding change in the reduced-form non-policy parameters should only occur if the Lucas critique does not hold (again, assuming policy impacts the relevant variable and, of course, that the model is reasonably well-specified). It is this possibility that we consider in our empirical analysis.

Table 3 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 48% to as high as 88%. 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 the probability of co-movement varying between 3% and 52%. 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 statistically significant effects on the unemployment rate throughout the sample period. So, if the Lucas critique holds, the reduced-form parameters should all move together.

[11] Although we need to consider changes in conditional mean parameters for the reduced-form VAR model, our approach does allow for identification of changes 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 3: Posterior median values of fractions of $k_{it} = k_{jt}, i, j = 1, 2, \cdots, 6, i \neq j$

$k_{it} = k_{jt}, i \neq j$	k_2	k_3	k_4	k_5	k_6
k_1	0.7733	0.1628	0.7791	0.4826	0.8547
	(0.0340)	(0.0224)	(0.0299)	(0.2125)	(0.0261)
k_2	–	0.1395	0.7907	0.4826	0.8721
	–	(0.0252)	(0.0335)	(0.2222)	(0.0302)
k_3	–	–	0.1395	0.5174	0.0349
	–	–	(0.0287)	(0.2741)	(0.0175)
k_4	–	–	–	0.4826	0.8779
	–	–	–	(0.2295)	(0.0308)
k_5	–	–	–	–	0.4826
	–	–	–	–	(0.2813)

Standard deviations are listed in parentheses.

Table 4: Test of the Lucas critique: Posterior probabilities of changes in parameter blocks conditional on a change in intercept or slope parameters in the interest rate equation

	Probability	95% Credible Interval
Inf[a]	0.8987[b]	[0.8809, 0.9069]
Unem	0.0297	[0.0280, 0.0300]
VarErr	0.9918	[0.9902, 0.9921]

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

As an even more direct way to look at co-movement related to changes in policy parameters, Table 4 reports on changes in blocks of parameters conditional on a change in intercept or slope parameters in 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 a 90% or higher probability of a change in the other parameters, except for the slopes in the unemployment equation, which only have a 3% conditional probability of change. However, looking back at Table 2, it is clear that certain parameters almost always change. For example, the structural shock variances appear to change about 98% of the time. So whenever the interest rate slope parameters change 88% of the time, we would expect a high conditional probability that the variance-covariance matrix changes too. But the question remains as to whether these parameter changes are causally related to each other.

To determine whether simultaneous parameter changes are merely coincidental, we calculate the correlation between changes in policy parameters and non-policy parameter blocks conditional on changes in the intercept or slope parameters from the interest rate equation. Table 5 reports posterior inferences for these correlations. Most of the correlations are essentially zero.[12] That is, even if policy and non-policy parameters change at the same time, they do not appear to change together in any systematic fashion. Thus, the evidence argues against the Lucas critique. Nonetheless, there are some statistically significant correlations, suggesting that the small correlations are not merely the consequence of the modeling assumption of independent switching in different blocks of parameters. For example, one of the more significant correlations is $corr(\Delta c_t^i, \Delta\overline{\theta}_{\pi t}^{\pi}) = 0.3162$, which suggests an increase in the intercept for the interest rate equation is positively related to higher persistence in inflation. The other two most significant correlations are $corr(\Delta\overline{\theta}_{\pi t}^{i}, \Delta\overline{\theta}_{ut}^{\pi}) = -0.3050$ and $corr(\Delta\overline{\theta}_{ut}^{i}, \Delta\overline{\theta}_{ut}^{\pi}) = -0.3540$, which suggest that a larger magnitude response of the interest rate to inflation or the unemployment rate (assuming a general positive response to inflation and a negative response to the unemployment rate) corresponds to a smaller tradeoff between the unemployment rate and inflation. These correlations suggest that systematic policy can affect private sector behaviour, but their reasonably small magnitude is a far cry from what one would expect if the persistence of inflation or changes in the slope of the Phillips curve were driven primarily by changes in the policy regime.

Figure 3 plots the posterior medians of first principal components for innovations to the policy and non-policy blocks, respectively. Both principal components explain well over 80% of the overall innovations in the corresponding blocks over time. It can be seen from the figure that variation in the policy block does not systematically relate to variation in the non-policy block (the sample correlation between the two series is only 0.1285). Notably, the principal component related to the policy block is reasonably persistent, while it is closer to white noise for the non-policy block. This finding supports the idea that changes in non-policy parameters are more likely driven by random shifts in technology and preferences than by changes in systematic monetary policy.

One alternative explanation for our findings could be model misspecification. A key assumption in our analysis is that we are capturing the main information set considered by policymakers when setting monetary policy. However, if this assumption is flawed, changes in policy parameters could be reflecting systematic responses to omitted variables. However, to the extent that inflation and the unemployment rate are the main drivers of systematic policy, as suggested by many variants of the Taylor rule, and assuming there have been shifts in this systematic policy throughout the postwar period, the Lucas critique clearly implies that we should find a stronger relationship between changes in policy parameters and the other reduced-form parameters than we actually do.[13]

[12] The results are robust for leads and lags of cross correlations, with posterior medians always almost exactly equal to zero. Furthermore, the results are robust to consideration of only ‘large” (greater than one standard deviation) changes in policy parameters.

[13] Another intriguing caveat to using our VAR model for testing the Lucas critique is that fiscal policy may have acted to offset the effects of changes in systematic monetary policy on the private sector. However, when we look at a measure of fiscal shocks based on Blanchard and Perotti’s (2002) constant-parameter VAR model, we find no

Table 5: Contemporaneous cross correlations between changes in policy parameters and non-policy parameters conditional on a change in the corresponding policy parameters in the interest rate equation

Non-Policy Block	Policy Block					
	Δc_t^i	$\Delta\overline{\theta}_{\pi t}^i$	$\Delta\overline{\theta}_{ut}^i$	$\Delta\overline{\theta}_{it}^i$	$\Delta a_{31,t}$	$\Delta a_{32,t}$
Δc_t^π	-0.0276	0.0189	0.0661	-0.0125	0.0031	0.0033
	(0.2046)	(0.1492)	(0.1424)	(0.1450)	(0.1503)	(0.1455)
Δc_t^u	0.1415	0.0111	-0.0409	0.0260	-0.0014	-0.0016
	(0.2007)	(0.1484)	(0.1418)	(0.1471)	(0.1507)	(0.1464)
$\Delta\overline{\theta}_{\pi t}^\pi$	0.3162	-0.1886	-0.0933	0.1467	-0.0010	0.0034
	(0.1305)	(0.1473)	(0.1443)	(0.1489)	(0.1488)	(0.1474)
$\Delta\overline{\theta}_{ut}^\pi$	-0.0739	-0.3050	-0.3540	-0.1559	0.0007	0.0036
	(0.1453)	(0.1386)	(0.1275)	(0.1470)	(0.1498)	(0.1526)
$\Delta\overline{\theta}_{it}^\pi$	-0.1847	-0.0344	-0.1519	-0.0273	-0.0018	0.0018
	(0.1380)	(0.1429)	(0.1362)	(0.1444)	(0.1502)	(0.1504)
$\Delta\overline{\theta}_{\pi t}^u$	-0.0475	-0.0027	-0.0069	-0.0228	-0.0005	0.0018
	(0.0845)	(0.0871)	(0.0856)	(0.0851)	(0.1497)	(0.1524)
$\Delta\overline{\theta}_{ut}^u$	0.0044	-0.0050	-0.0064	-0.0021	0.0008	0.0018
	(0.0858)	(0.0859)	(0.0860)	(0.0853)	(0.1505)	(0.1495)
$\Delta\overline{\theta}_{it}^u$	-0.0123	-0.0133	-0.0365	-0.0518	-0.0021	0.0022
	(0.0866)	(0.0847)	(0.0835)	(0.0832)	(0.1513)	(0.1506)
$\Delta a_{21,t}$	-0.0014	-0.0023	-0.0009	-0.0018	-0.0003	0.0009
	(0.0837)	(0.0819)	(0.0825)	(0.0823)	(0.1444)	(0.1513)

Standard deviations are listed in parentheses. Δ is the difference operator. For the equation for variable k, c_t^k is the intercept and $\overline{\theta}_{jt}^k$ is the sum of slopes on variable j, where $j, k = \pi, u, i$, corresponding to inflation, the unemployment rate, and the interest rate, respectively.

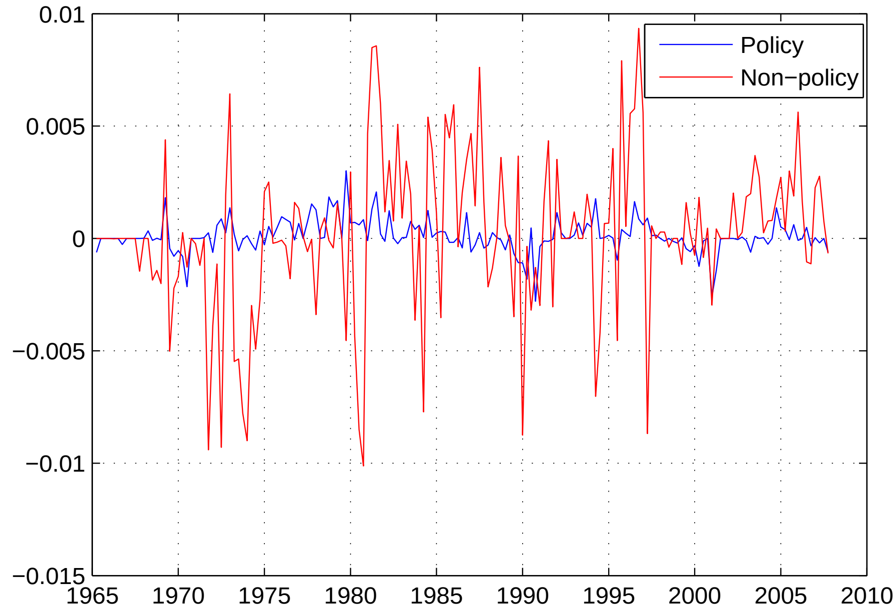


Figure 3: Posterior medians of first principal components for innovations to policy and non-policy blocks.

6 Evolution of Impulse Responses

There is considerable academic debate about whether monetary policy is responsible for stabilizing the U.S. 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.[14]

Figures 4-6 plot the evolution of impulse responses for inflation, the unemployment rate, and the interest rate to a one percentage point monetary policy shock at selected dates: 1975Q1, 1981Q3, 1996Q1 and 2006Q3.[15] The estimated magnitudes of the responses of inflation, the

relationship between the estimated fiscal shocks from that model, which should reflect omitted shift in systematic fiscal policy, and the changes in parameters from the interest rate equation in our model. We leave analysis of a mixture innovations fiscal VAR model for future research.

[14]Another 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.

[15]For easy comparison, these dates are the same as those considered in Primiceri (2005) or Koop et al. (2009). The first date is an NBER trough, the second date is an NBER peak, and the last two dates are more normal times. Thus,

unemployment rate, and the interest rate are 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 contrast somewhat with those in Primiceri (2005), Koop et al. (2009) and Boivin and Giannoni (2006).[16] Boivin and Giannoni (2006) study time-invariant VAR models of inflation, output, and the interest rate for two subsamples 1959Q1 - 1979Q3 and 1979Q4 - 2002Q2 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 statistical test of whether impulse responses have 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 are 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 2, the BEQ model receives the strongest 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.

---

we can see how responses change over the business cycle, as well as how they have evolved in recent times at a similar stage of the business cycle. Note that we consider the Fed Funds Rate and inflation based on the PCE deflator, while the previous studies consider the 3-month Treasury bill rate as a proxy for the policy instrument and measure inflation using the GDP deflator. However, we find that the impulse response results are generally robust to considering other measures of the interest rate and inflation.

[16]There is an apparent “price puzzle” in 1975Q1 and 1981Q3, which is common for small monetary VAR models with triangular identification schemes for pre-1980 U.S. 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.

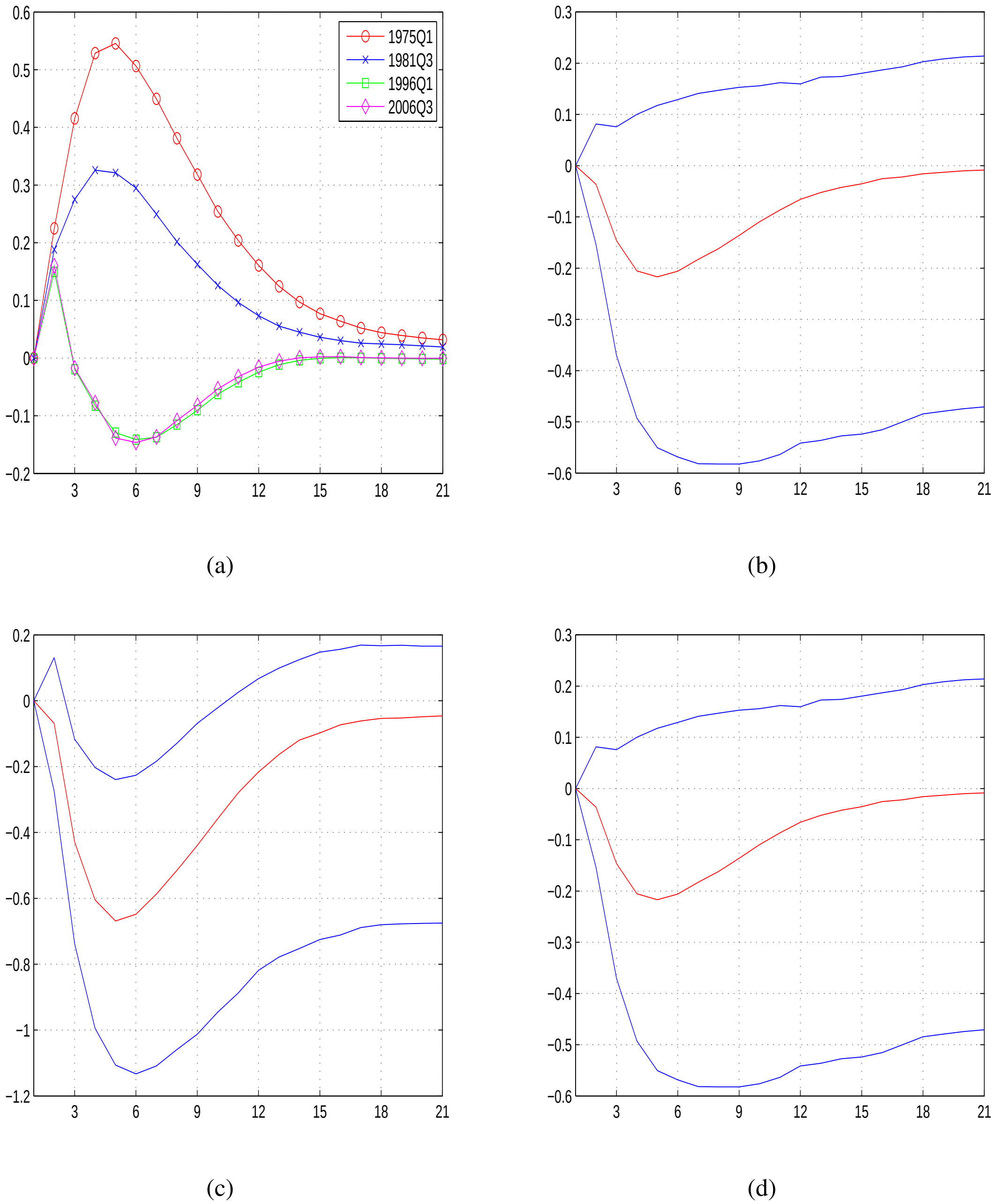


Figure 4: Response of inflation to a one percentage point monetary policy shock in 1975Q1, 1981Q3, 1996Q1 and 2006Q3: (a) medians of impulse responses; (b) response in 1981Q3 minus response in 1975Q1 with 80% equal-tailed credible interval; (c) response in 1996Q1 minus response in 1975Q1 with 80% equal-tailed credible interval; (d) response in 2006Q3 minus response in 1996Q1 with 80% equal-tailed credible interval.

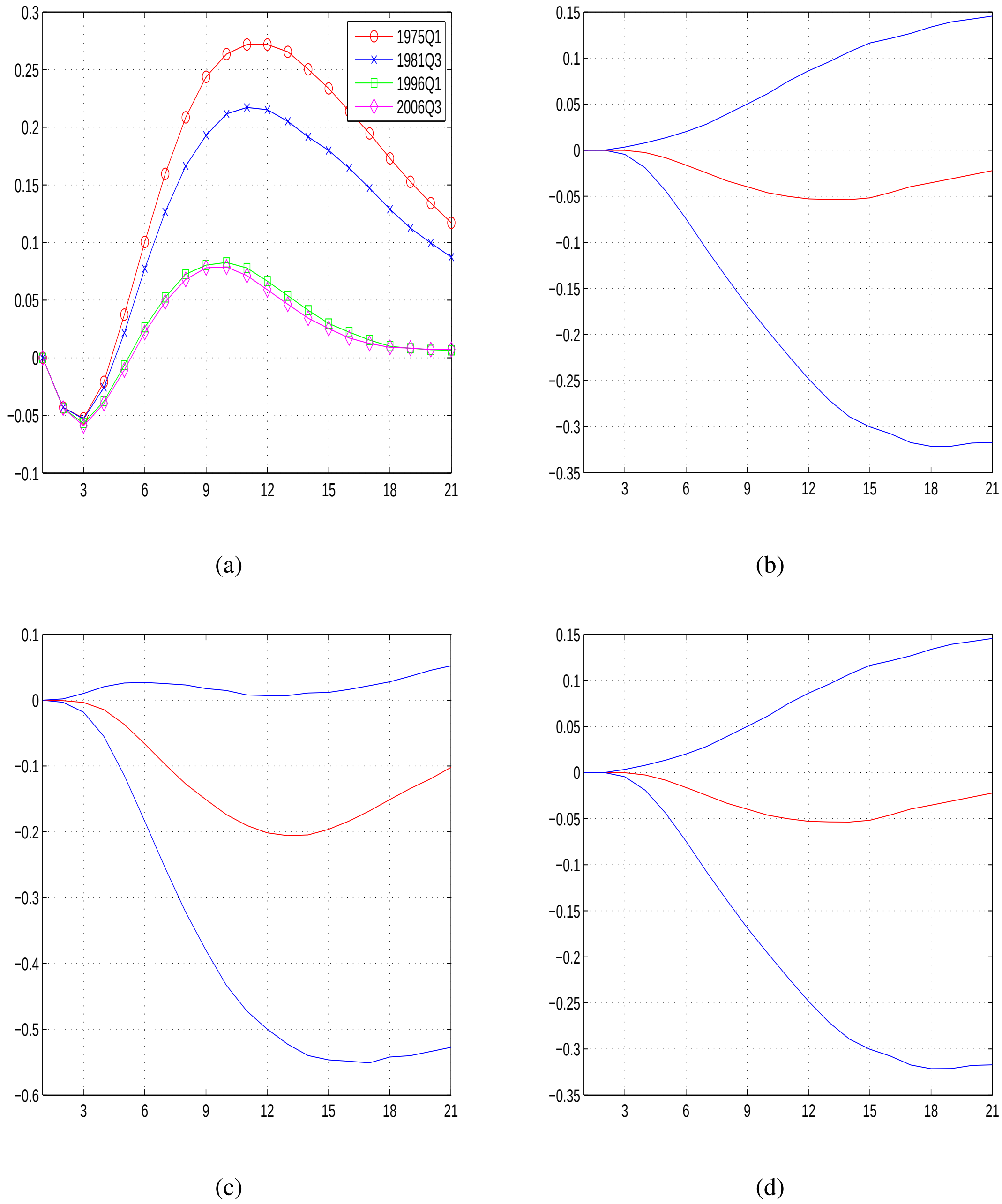
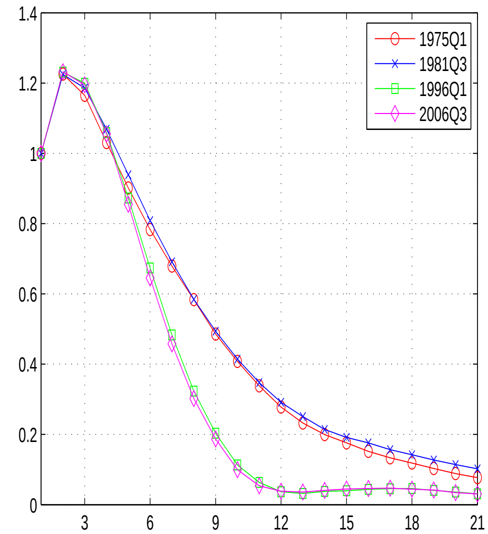
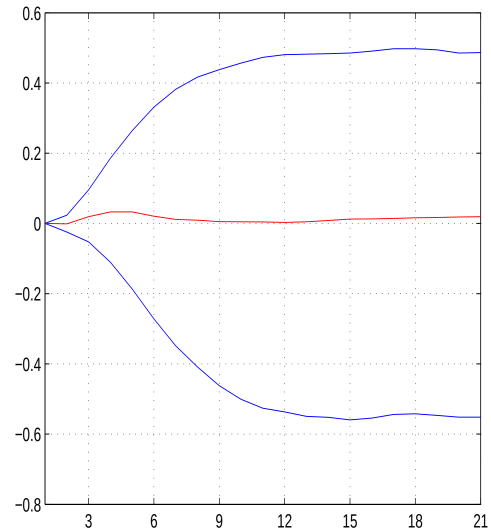


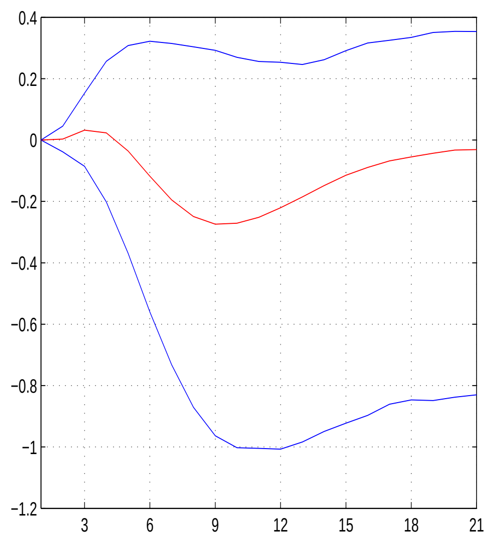
Figure 5: Response of the unemployment rate to a one percentage point monetary policy shock for 1975Q1, 1981Q3, 1996Q1 and 2006Q3: (a) medians of impulse responses; (b) response in 1981Q3 minus response in 1975Q1 with 80% equal-tailed credible interval; (c) response in 1996Q1 minus response in 1975Q1 with 80% equal-tailed credible interval; (d) response in 2006Q3 minus response in 1996Q1 with 80% equal-tailed credible interval.



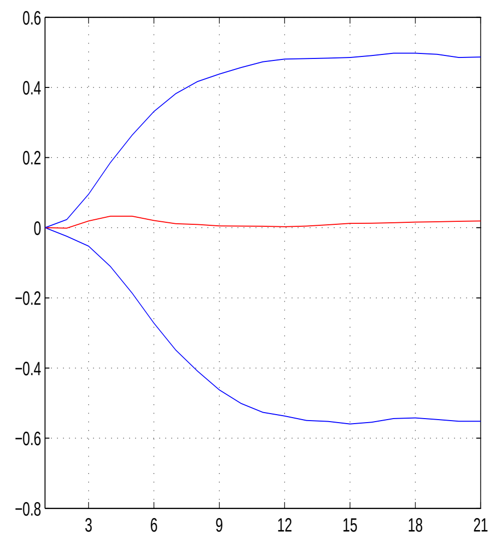
(a)



(b)



(c)



(d)

Figure 6: Response of the interest rate to one percentage point monetary policy shock for 1975Q1, 1981Q3, 1996Q1 and 2006Q3: (a) medians of impulse responses; (b) response in 1981Q3 minus response in 1975Q1 with 80% equal-tailed credible interval; (c) response in 1996Q1 minus response in 1975Q1 with 80% equal-tailed credible interval; (d) response in 2006Q3 minus response in 1996Q1 with 80% equal-tailed credible interval.

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-accelarating 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, allow for a few discrete jumps at certain periods of time (e.g., Papell et al., 2000), force the NRU to be a function of time using a "spline" (e.g., Staiger et al., 1997), or employ other techniques such as calibrated unobserved-components models (e.g., Gordon, 1997), low-pass filtering (e.g., Staiger et al., 2001) and the Hodrick-Prescott filter (e.g., Ball and Mankiw, 2002). King and Morley (2007) endogenize the NRU as the steady-state derived from a VAR model in the spirit of the following quote from Phelps (1994):

> *"In a useful shorthand one may characterize the theory here as endogenizing the natural rate of unemployment – 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 the unemployment rate following Beveridge and Nelson (1981) by calculating the long-run forecast in levels $\overline{y}_t = \lim_{h\to\infty} E_t y_{t+h}$ conditional on the information set available at time t.

King and Morley (2007) assume a time-invariant VAR model for output growth, inflation, and the first difference of the unemployment rate. However, as shown above, there is considerable variation in the parameters for the VAR models of inflation, the unemployment rate, and the interest rate considered in this paper. Thus, it would be interesting to see if this time variation has any implications for the time-varying NRU. To do this, we follow King and Morley (2007) by first casting a VAR model 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 constant at their current values as time goes forward—i.e., in each period of time, a time-invariant VAR model is assumed based on the time-varying parameter estimates for that period.[17] Conditional

[17] This "local-to-date" assumption is common in the literature on bounded rationality and learning (see the "anticipated-utility" model in Kreps, 1998). It should be noted that when we simulate a long-run forecast of the

on the information set available at time t, including g_t and F_t, the long-run forecast is

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

where s_u is a selector vector for the unemployment rate and $\overline{u}_t$ is the time-varying NRU.[18]

The natural rate of unemployment and the actual unemployment rate are plotted in Figure 7. The point estimates for the NRU range from 4.0 − 7.2 percent, which is less volatile compared to the range of 1.8 − 9.5 percent for the point estimates obtained by King and Morley (2007), but is comparable to Phelps's (1994) estimates. The uncertainty around 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. In addition to the difference in the range of the NRU, our estimate is a lot smoother than that reported in King and Morley (2007). This is possibly due to the fact that our model allows blocks of the VAR coefficients to stay constant at their previous values in some periods, so that the trend of the unemployment rate is not forced to be stochastic in every period.

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.[19] Following Gordon (1997), we investigate a "triangle" model (although without explicit supply shocks). Also, Peach et al. (2011) recently consider a threshold Phillips curve model with respect to cyclical unemployment to investigate possible nonlinearities. They find that the tradeoff between inflation and cyclical unemployment depends on slack of the labor market. Related to this, we sometimes include interaction terms to capture possible state dependence in the tradeoff between inflation and cyclical unemployment. Specifically, we consider the following four specifications:

$$\text{Linear:} \quad \pi_t = \sum_{i=1}^{4} \delta_i \pi_{t-i} + \beta_1 u_t^C + \omega_t, \tag{7}$$

$$\text{NL-Full:} \quad \pi_t = \sum_{i=1}^{4} \delta_i \pi_{t-i} + \beta_1 u_t^C + \beta_2 u_t \cdot u_t^C + \beta_3 \pi_{t-1} \cdot u_t^C + \omega_t, \tag{8}$$

unemployment rate based on all shocks in the model, including those to the time-varying parameters, we obtain a similar measure of the NRU. For consistency with our impulse response analysis, for which a simulated approach would be much more computationally demanding, we report the results for the local-to-date measure.

[18]Although we do not impose stationarity on F_t, it turns out that almost all of the draws from our sampler satisfy the stationary conditions, making the NRU well behaved. Note that King and Morley (2007) calculate the long-run forecast of the level of the unemployment rate given a stationary VAR model that includes its first differences. By contrast, we calculate the long-run forecast of the level of the unemployment rate given a VAR model that includes its levels. However, the time-varying intercept allows for a stochastic trend in the unemployment rate, so the two approaches are similar as long as the companion matrix F_t for our VAR model has all of its eigenvalues less than one in modulus.

[19]We construct cyclical unemployment by subtracting the median NRU from the actual unemployment rate.

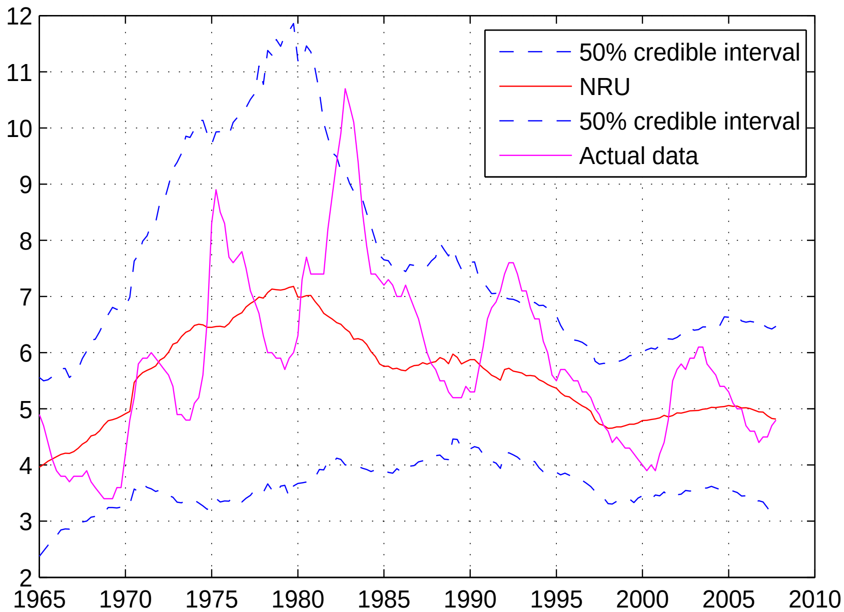


Figure 7: Natural rate of unemployment: posterior median and 50% equal-tailed credible interval

$$\text{NL-}\pi\text{:} \quad \pi_t = \sum_{i=1}^{4} \delta_i \pi_{t-i} + \beta_1 u_t^C + \beta_3 \pi_{t-1} \cdot u_t^C + \omega_t, \tag{9}$$

$$\text{NL-}u\text{:} \quad \pi_t = \sum_{i=1}^{4} \delta_i \pi_{t-i} + \beta_1 u_t^C + \beta_2 u_t \cdot u_t^C + \omega_t, \tag{10}$$

where u_t^C is our measure of cyclical unemployment and ω_t is a regression error. "Linear" denotes the linear model. "NL-Full" denotes a nonlinear model with interaction terms $\pi_{t-1} \cdot u_t^C$ and $u_t \cdot u_t^C$. "NL-π" and "NL-u" denote nonlinear models related only to the level of lagged inflation or the unemployment rate, respectively.

Table 6 reports regression results based on OLS for the full sample and two subsamples of 1965Q1-1990Q4 and 1991Q1-2007Q4.[20] Two 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. Regardless, the 95% confidence intervals for the sum of the δ's reported in Table 6 always contain 1. Hence, the natural rate

[20] Our results are robust to also including an intercept or lags of cyclical unemployment. Also, in terms of a possible error-in-variables problem due to measurement error for cyclical unemployment, we have conducted a Hausman test using the lagged unemployment rate and lagged first differences of the unemployment rate as instruments and found no evidence of endogeneity. Meanwhile, the timing of the subsamples is chosen based on Atkeson and Ohanian (2001) and King and Morley (2007).

hypothesis is supported by the data.[21] Second, there is strong evidence supporting the existence and nonlinearity of a short-run tradeoff between inflation and cyclical unemployment for the full sample and the first subsample. For the full sample, the best specification is the nonlinear model related to the level of lagged inflation only (NL-π) according to adjusted R-squared, AIC and SIC. The nonlinearity coefficient $\beta_3 = -0.134$ is statistically significant at the 5% level, implying that a higher level of inflation corresponds to a greater tradeoff. A likely explanation for this nonlinearity is that inflation expectations are not as well-anchored at high levels of inflation, with cyclical unemployment generating a larger change in inflation in this case. The same result and general reasoning apply for the first subsample in which the best model is the full nonlinear model (NL-Full) measured by adjusted R-squared and AIC, or the NL-π model based on the SIC. In this case, the nonlinearity coefficient β_3 is statistically significant at 5% level and quantitatively similar for both models, specifically, -0.145 in the NL-Full model and -0.143 in the NL-π model. By contrast, for the second subsample, the short-run tradeoff has weakened and possibly even disappeared, as none of the tradeoff coefficients β_1,β_2 and β_3 is statistically significant for the models considered. The vanishing tradeoff between inflation and cyclical unemployment in the second subsample is in accordance with the findings of Atkeson and Ohanian (2001) and could reflect a strong anchoring of inflation expectations in the recent sample period (see IMF, 2013).

In addition to the regression analysis, we investigate time variation in the short-run tradeoff between inflation and cyclical unemployment using the impulse responses discussed in the previous section. Specifically, we consider the ratio of the 3-9 quarter average response of inflation relative to the 3-9 quarter average response of the unemployment rate for each structural shock. Figure 8 plots the posterior medians of the ratios of the inflation and unemployment rate responses for each structural shock. The short-run tradeoffs 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 50% equal-tailed credible intervals, as reported in Figure 9.[22] This tradeoff strengthened until around 1977 and then weakened and possibly disappeared by the mid-1990s, consistent with our findings on a nonlinear Phillips curve with respect to the level of lagged inflation based on the OLS regressions reported in Table 6 given that inflation was relatively high in the 1970s.

[21] The 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).

[22] To the extent that the structural shocks in the unemployment rate and interest rate equations represent aggregate demand shocks, we would expect a similar tradeoff for both shocks. So the differences in the estimated tradeoffs could reflect estimation uncertainty, with the effects of monetary policy shocks not as well identified given their relative unimportance in explaining fluctuations in the unemployment rate and inflation. This is consistent with the relatively wide posterior bands for the tradeoff given a monetary policy shock compared to an unemployment shock. Meanwhile, a common pattern of a diminished tradeoff for both shocks is clearly evident in Figure 8.

Table 6: Phillips curve regression results: OLS estimates

	Full Sample				1965Q1 - 1990Q4				1991Q1 - 2007Q4			
Regressors: Coeff.	Linear	NL-Full	NL-π	NL-u	Linear	NL-Full	NL-π	NL-u	Linear	NL-Full	NL-π	NL-u
$\pi_{t-1} : \delta_1$	0.571	0.546	0.547	0.571	0.595	0.533	0.582	0.590	0.348	0.351	0.351	0.348
	(0.077)	(0.073)	(0.073)	(0.077)	(0.099)	(0.094)	(0.094)	(0.099)	(0.127)	(0.129)	(0.127)	(0.128)
$\pi_{t-2} : \delta_2$	0.079	0.107	0.108	0.079	0.055	0.128	0.093	0.054	0.202	0.201	0.201	0.202
	(0.088)	(0.084)	(0.084)	(0.089)	(0.116)	(0.110)	(0.110)	(0.116)	(0.131)	(0.133)	(0.132)	(0.132)
$\pi_{t-3} : \delta_3$	0.158	0.151	0.151	0.158	0.125	0.164	0.111	0.125	0.348	0.348	0.348	0.348
	(0.088)	(0.084)	(0.084)	(0.089)	(0.116)	(0.109)	(0.110)	(0.116)	(0.127)	(0.129)	(0.128)	(0.128)
$\pi_{t-4} : \delta_4$	0.196	0.196	0.200	0.194	0.231	0.187	0.225	0.226	0.069	0.065	0.067	0.068
	(0.078)	(0.074)	(0.074)	(0.078)	(0.099)	(0.094)	(0.094)	(0.099)	(0.143)	(0.146)	(0.143)	(0.145)
$u_t^C : \beta_1$	-0.390	0.074	0.238	-0.471	-0.457	0.059	0.298	-0.810	-0.082	-0.022	0.026	-0.110
	(0.082)	(0.313)	(0.161)	(0.305)	(0.109)	(0.457)	(0.245)	(0.421)	(0.119)	(0.722)	(0.396)	(0.650)
$u_t \cdot u_t^C : \beta_2$	–	0.023	–	0.011	–	0.036	–	0.044	–	0.008	–	0.005
	–	(0.036)	–	(0.039)	–	(0.048)	–	(0.051)	–	(0.105)	–	(0.103)
$\pi_{t-1} \cdot u_t^C : \beta_3$	–	-0.136	-0.134	–	–	-0.145	-0.143	–	–	-0.048	-0.046	–
	–	(0.030)	(0.030)	–	–	(0.039)	(0.042)	–	–	(0.164)	(0.162)	–
δ^*	1.004	1.001	1.006	1.002	1.006	1.013	1.010	0.995	0.966	0.965	0.967	0.965
	(0.166)	(0.158)	(0.157)	(0.160)	(0.216)	(0.204)	(0.205)	(0.216)	(0.264)	(0.269)	(0.266)	(0.267)
Adjusted R^2	0.809	0.828	0.829	0.808	0.731	0.764	0.758	0.731	0.319	0.298	0.309	0.308
AIC	3.051	2.956	2.947	3.063	3.406	3.297	3.310	3.418	2.125	2.182	2.153	2.154
BIC/SIC	3.144	3.087	3.058	3.174	3.537	3.479	3.466	3.575	2.288	2.411	2.349	2.350
Durbin-Watson stat	1.964	2.085	2.078	1.867	1.962	2.013	2.088	1.977	1.922	1.922	1.922	1.922

Standard deviations are listed in parentheses. The sum of the δ's is given by $\delta^* = \sum \delta_j,\ j = 1, .., 4$.

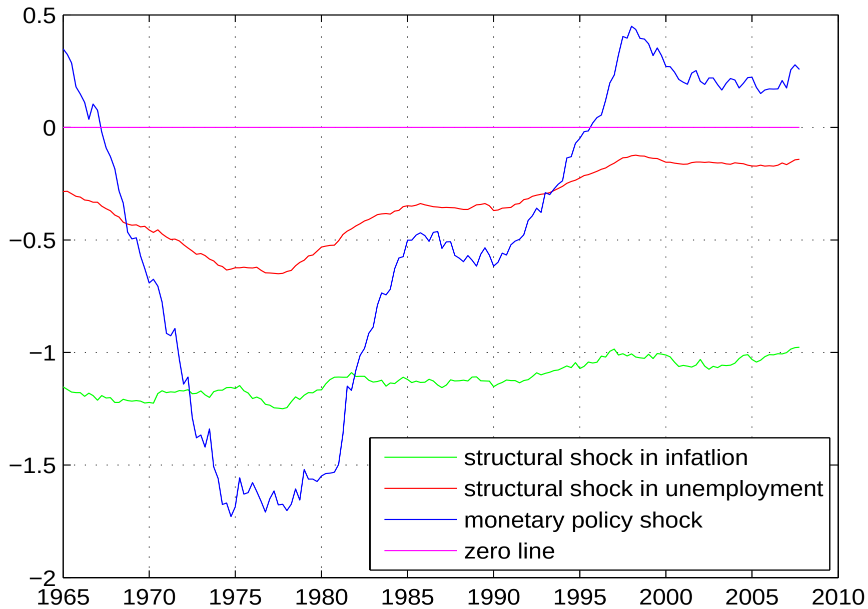


Figure 8: The posterior medians of the ratios of the inflation and unemployment rate responses (averaged over the 3-9 quarter horizon) for each structural shock

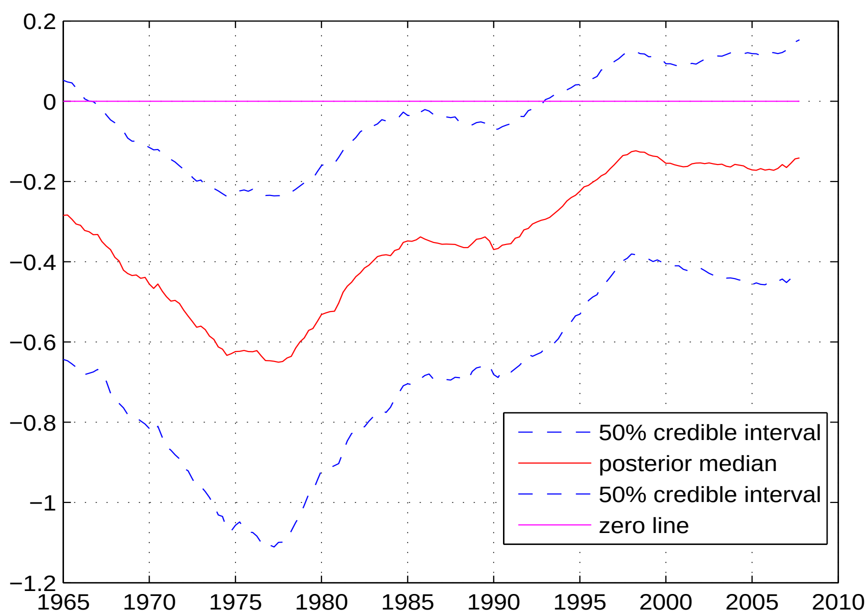


Figure 9: The posterior median and 50% equal-tailed credible interval for the ratio of the inflation and unemployment rate responses (averaged over the 3-9 quarter horizon) for a structural shock to the unemployment rate

Figure 10 presents a related measure of the decline in the short-run tradeoff 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 1975Q1 vs. 1990Q3, 1975Q1 vs. 1991Q1, 1975Q1 vs. 2000Q1 and 1990Q3 vs. 2000Q1. 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 10 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 $F_{1975}^{1990} = 85.2\%$. This means that 85.2% of the simulated ratios in 1990Q3 are greater than those in 1975Q1. The equivalent statistics for the other dates are $F_{1975}^{1991} = 85\%, F_{1975}^{2000} = 91.8\%$, and $F_{1990}^{2000} = 84.7\%$. Thus, from Figure 10 and the corresponding F_A^B statistic, we can conclude that the short-run tradeoff between inflation and cyclical unemployment has declined since the beginning of the 1990s.

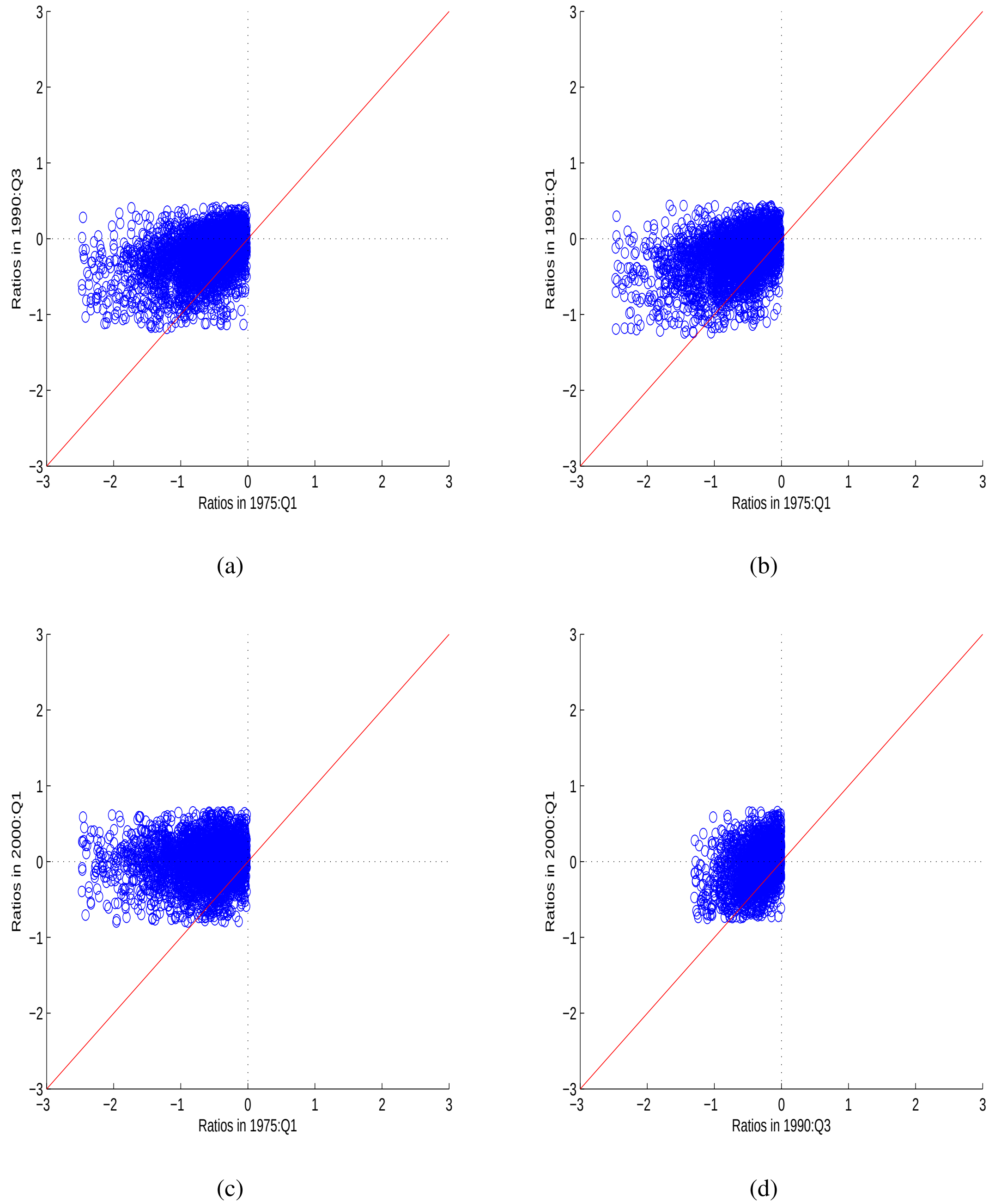


Figure 10: 95% joint credible sets of ratios of inflation and unemployment rate responses (averaged over the 3-9 quarter horizon) for a structural shock to the unemployment rate across certain periods of time: (a) 1975Q1 vs. 1990Q3; (b) 1975Q1 vs. 1991Q1; (c) 1975Q1 vs. 2000Q1; (d) 1990Q3 vs. 2000Q1. 95% joint credible sets are constructed by excluding 2.5% equal-taileded 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. Notably, 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 part of the flexible variation allowed in the 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 and reject the empirical relevance of Lucas critique notion that changes in policy parameters drive changes in reduced-form parameters.

Even though we do not find support for the Lucas critique, our estimates suggest that the structure of the U.S. economy has evolved considerably over the postwar period, with diminished effects of monetary policy shocks on inflation and changes in the slope of the Phillips curve in recent years. However, it is notable that the structural changes have been gradual and small enough in the recent sample that estimates from the model could be useful for predicting the effects of monetary policy in the future, at least once the U.S. economy exits the zero-lower-bound period.

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.

Blanchard, O. and R. Perotti (2002), "An empirical characterization of the dynamic effects of changes in government spending and taxes on output", *Quarterly Journal of economics*, Vol. 117, 1329-1368.

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 U.S.", *American Economic Journal: Macroeconomics*, Vol. 2, 43 - 69.

Cogley, T. and T. Sargent (2001), "Evolving Post-World War II U.S. 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 U.S.", *Review of Economic Dynamics*, Vol. 8, 262 - 302.

Collard, F. and F. Langot (2001), "Structural Inference and The Lucas Critique", *Annales d'economie 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.

Giordani, P. and R. Kohn (2008), “Efficient Bayesian inference for multiple change-point and mixture innovation models”, *Journal of Business and Economic Statistics*, Vol. 26, 66-77.

Giordani, P. and M. Villani (2010), “Forecasting macroeconomic time series with locally adaptive signal extraction”, *International Journal of Forecasting*, Vol. 26, 312-325.

Gordon, R. (1997), “The Time-varying NAIRU and Its Implications for Economic Policy”, *Journal of Economic Perspectives*, Vol. 11, 11 - 32.

Groen, J. J., R. Paap and F. Ravazzolo (2013), “Real-time inflation forecasting in a changing world”, *Journal of Business and Economic Statistics*, Vol. 31, 29-44.

IMF (2013), “The dog that didn’t bark: Has inflation been muzzled or was it just sleeping?”, In *World Economic Outlook, April*, 79 - 96. Washington, DC: International Monetary Fund.

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.

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.

Koop, G. and S. Potter (2007). “Estimation and forecasting in models with multiple breaks”, *The Review of Economic Studies*, Vol. 74, 763-789.

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.

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.

Lindé, 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 Keyensian 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.

Peach, R. W., R. W. Rich and Cororaton, A. (2011). “How does slack influence inflation?”, *Current Issues in Economics and Finance*, Vol. 17, No. 3.

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.

Primiceri, G. and M. Del Negro (2013), “Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum”, manuscript.

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 U.S. 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 U.S. 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, $\lambda = \left\{\lambda_{1j}, \lambda_{2j}\right\}_{j=1}^{6}$ and $K^T = \left(k_1^T, k_2^T, k_3^T, k_4^T\right)$, 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), as follows:

A1. Drawing latent variables K^T, k_5^T and k_6^T

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

$$y_t = X_t'\theta_t + A_t^{-1}\varepsilon_t, \tag{A.1}$$

$$\theta_t = \theta_{t-1} + K_t\xi_t. \tag{A.2}$$

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

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

where $K_{\backslash t} = K^T \backslash K_t$ and $x^{s,t} = (x_s, x_{s+1}, \cdots, x_t)$, $s < t$. 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):

$$\hat{y}_t = D_t\alpha_t + \Sigma_t\epsilon_t, \tag{A.3}$$

$$\alpha_t = \alpha_{t-1} + k_{5t}\eta_t, \tag{A.4}$$

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), we consider the non-Gaussian linear state-space model with respect to the state $h_t = \ln \sigma_t$ as follows:

$$y_t^{**} = 2h_t + e_t, \tag{A.5}$$

$$h_t = h_{t-1} + k_{6t}\zeta_t, \tag{A.6}$$

where $y_t^{**} = [y_{1t}^{**} \quad y_{2t}^{**} \quad y_{3t}^{**}]'$, $y_t^{**} = \ln[(y_t^*)^2 + c]$, $y_t^* = A_t(y_t - X_t'\theta_t)$, $c = 0.001$ and $e_t = [e_{1t} \quad e_{2t} \quad e_{3t}]'$ 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 A7 below.

A2. Drawing parameters of Beta priors λ

Denote the Beta priors for p_j, $j = 1, 2, \cdots, 6$ by $Beta(\underline{\lambda}_{1j}, \underline{\lambda}_{2j})$, therefore, the posterior distribution of p_j is $Beta(\overline{\lambda}_{1j}, \overline{\lambda}_{2j})$, where [23]

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

A3. Drawing reduced VAR parameters θ^T

[23] Hereafter, the underlined parameters stand for the parameters of priors and the overlined parameters represent the parameters of posteriors.

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 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

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

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 $P_{t|t+1}, t = 1, 2, \cdots, T-1$, 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 $\theta_t, t = 1, 2, \cdots, T-1$. Please see the details of the Gibbs sampling procedure in Appendix B.

A4. Drawing hyperparameter Q

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

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

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

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

where

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

A5. Drawing covariances α^T

Again, consider the Gaussian linear state-space model A.3 and A.4 under the assumption of block-diagonal S. Because $\hat{y}_{1,t}$ is determined by exogenous identity shock ϵ_{1t} and $\sigma_{11,t}$, then conditional on other blocks, $\hat{y}_{1,t}$ is predetermined in $\hat{y}_{2,t}$'s equation. Similarly, $\hat{y}_{1,t}$ and $\hat{y}_{2,t}$ are predetermined in $\hat{y}_{3,t}$'s equation. Therefore, α_t can be obtained by applying the 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

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

A6. 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, $j = 1, 2$, is as follows:

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

where

$$\overline{S}_j^{-1} = \left[\underline{S}_j^{-1} + \sum_{t=1}^{T} (\alpha_{j,t+1} - \alpha_{j,t})(\alpha_{j,t+1} - \alpha_{j,t})'\right]^{-1} \quad and \quad \overline{\nu}_{S_j} = \underline{\nu}_{S_j} + \sum_{t=1}^{T} k_{5t}.$$

A7. Drawing stochastic volatility σ^T

The stochastic volatility σ^T is drawn from the non-Gaussian linear state-space model A.5 and A.6 based on a mixture of seven normals approximation as in Kim et al. (1998) with component probabilities q_l, means $m_l - 1.2704$ and variances $v_l^2, l = 1, 2, \cdots, 7$. 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, we can employ the same mixture of normals to approximate any element in e_t.

Define the state-indicator matrix $s^T = [s_1, s_2, \cdots, s_T]'$, $s_t = [s_{1t}, s_{2t}, s_{3t}]'$, $s_{jt} \in \{1, 2, \cdots, 7\}$, $j = 1, 2, 3$ and $t = 1, 2, \cdots, T$, which shows in each period of time which member of the mixture of normals is used for each element of e_t. Then, s^T can be updated as in Kim et al. (1998) for each s_{jt} independently from the discrete density

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

where $f_N(\cdot)$ stands for the normal density.

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 standard Kalman filtering and backward recursions as in previous steps. Specifically, smoothed estimates of h_t can be drawn recursively from

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

Finally, the smoothed estimate of σ_t can be recovered by the transformation $\sigma_t = exp\{0.5h_t\}$.

A8. Drawing hyperparameter W

Note that $W \sim \mathbb{IW}(\underline{W}, \underline{\nu}_W)$, i.e. $W^{-1} \sim \mathbb{W}(\underline{W}^{-1}, \underline{\nu}_W)$, 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:

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

where

$$\overline{W}^{-1} = \left[\underline{W}^{-1} + \sum_{t=1}^{T} (h_{t+1} - h_t)(h_{t+1} - h_t)' \right]^{-1} \quad \textbf{and} \quad \overline{\nu}_W = \underline{\nu}_W + \sum_{t=1}^{T} k_{6t}.$$

Appendix B. Gibbs Sampling for State-Space Models

We cast the Gaussian linear state-space models considered 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

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

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, $\beta_{T-2}, \beta_{T-3}, \cdots, \beta_1$ are drawn from $N(\beta_{T-2|T-1}, P_{T-2|T-1}), N(\beta_{T-3|T-2}, P_{T-3|T-2}), \cdots, N(\beta_{1|2}, P_{1|2})$, respectively.