ONLINE APPENDIX
Time-Varying Effects of Oil Supply Shocks on the U.S. Economy

Christiane Baumeister
Bank of Canada
cbaumeister@bankofcanada.ca

Gert Peersman
Ghent University
gert.peersman@ugent.be

December 2011

Here we provide a detailed description of the data sources and of the construction of the data series. We outline the estimation of the time-varying VAR model including the choice of the priors and the steps of the Metropolis-within-Gibbs sampling algorithm to approximate the joint posterior distribution. We explain the computation of the impulse responses and the implementation of the sign restrictions to identify exogenous oil supply shocks. We describe the setup and the results of a Monte Carlo simulation exercise that shows that our econometric model is able to capture abrupt changes in a satisfactory manner. We report on the significance of time variation in the impact responses of the four endogenous variables. We present evidence that the main findings of the paper are robust to changes in the variables included in the model, to alternative model specifications, and to different identification assumptions.

A Data description

Monthly world oil production data measured in thousands of barrels of oil per day were obtained from the U.S. Energy Information Administration's (EIA) *Monthly Energy Review* starting in January 1973. Monthly data for global production of crude oil for the period $1953M4$ to $1972M12$ were taken from the weekly *Oil & Gas Journal* (issue of the first week

of each month). For the period $1947M1$ to $1953M3$, monthly data were constructed by interpolation of yearly world oil production data by means of the Litterman (1983) methodology using U.S. monthly oil production data from the EIA as an indicator variable.[1] Annual oil production data were obtained from *World Petroleum* $(1947-1954)$, the *Oil & Gas Journal* (end-of-year issues, $1954-1960$) and the EIA's *Annual Energy Review* $(1960-2010)$. Consistency between these different data sources was checked at each of the overlapping periods. Quarterly data are averages of monthly observations.

The nominal U.S. refiners' acquisition cost of imported crude oil was taken from the *Monthly Energy Review*.[2] Since this series is only available from January 1974 onwards, it was backcast until $1947Q1$ with the quarterly growth rate of the producer price index (PPI) for crude oil retrieved from the Bureau of Labor Statistics (BLS) database (WPU0561). Data were converted to quarterly frequency before backcasting by averaging over months. For the robustness checks with regard to the choice of the oil price variable, we use the quarterly average of the West Texas Intermediate (WTI) spot oil price obtained from the Federal Reserve Economic Data (FRED) database maintained by the St. Louis FED (OILPRICE) and of the nominal U.S. refiners' acquisition cost of composite[3] crude oil from the *Monthly Energy Review*. The latter was adjusted for price controls on domestic oil production for the period $1971Q3$ to $1974Q1$ as described in Mork (1989) and reconstructed backwards to $1947Q1$ in the same way as the imported refiners' acquisition cost series.

Quarterly seasonally adjusted series for U.S. real GDP (GDPC96: real gross domestic product, billions of chained 2005 dollars) and for the U.S. GDP deflator (GDPDEF: gross domestic product implicit price deflator) were obtained from the FRED database. Monthly seasonally adjusted data for U.S. industrial production (INDPRO: industrial production index, index $2007 = 100$), for the U.S. consumer price index (CPIAUCSL: consumer price index for all urban consumers: all items, index $1982-1984 = 100$), for the civilian unemployment rate, 16 years and older, seasonally adjusted (UNRATE), and for the effective

[1] Since this part of the data is only needed for the training sample to calibrate the priors based on the estimation of a fixed-coefficient VAR, the use of interpolated data as opposed to actual ones is of minor consequence.

[2] The refiners' acquisition cost of crude oil imports (IRAC) is a volume-weighted average price of all kinds of crude oil imported into the U.S. over a specified period. Since the U.S. imports more types of crude oil than any other country, it may represent the best proxy for a true "world oil price" among all published crude oil prices. The IRAC is also similar to the OPEC basket price.

[3] The entitlement system in force during the 1970s in the U.S. required buyers to purchase foreign and domestic oil in fixed proportions so that the aggregate price was a weighted average of these two kinds of oil.

federal funds rate (FEDFUNDS) were taken from the FRED database, where the latter three were converted to quarterly frequency by taking averages.

Data on annual energy consumption by sector measured in billions of British Thermal Units (BTU) were obtained from the *Annual Energy Review* (Tables $2.1c - 2.1f$) for the period $1974 - 2010$. Primary and secondary uses of petroleum in the industrial, commercial and part of the transportation sector (buses and heavy trucks) were aggregated to obtain a measure of total petroleum consumption in the production process. Data on highway transportation energy consumption by mode ($1970 - 2009$) were taken from the *Transportation Energy Data Book* (Table 2.7).

Annual data on global spare capacity of oil production for the period $1974 - 2010$ were taken from the IMF *World Economic Outlook* and updated with the EIA's *Short-Term Energy Outlook*. Spare capacity refers to production capacity that can be brought online within 30 days and sustained for 90 days. Global capacity utilization rates are calculated as a percentage of total potential annual world oil production, which is the sum of available spare capacity and actual oil production taken from the *Annual Energy Review*.

B Bayesian estimation of a VAR with time-varying parameters and stochastic volatility

Consider the time-varying vector autoregression model with stochastic volatility described by equations (1) to (7) in the main text.

B.1 Prior distributions and starting values

The priors for the initial states of the drifting coefficients, the covariances and the log volatilities, $p(\theta_0)$, $p(\alpha_0)$ and $p(\ln h_0)$ respectively, are assumed to be normally distributed, independent of each other and independent of the hyperparameters, which are the elements of Q, S and the σ_i^2 for $i = 1, ..., 4$. The priors are calibrated on the point estimates of a constant-coefficient VAR(4) estimated over the period $1947Q2 - 1972Q2$.

The unrestricted prior for the VAR coefficients is set to

$$\theta_0 \sim N\left[\widehat{\theta}_{OLS}, \widehat{P}_{OLS}\right] \tag{1}$$

where $\widehat{\theta}_{OLS}$ corresponds to the OLS point estimates of the training sample and $\widehat{P}_{OLS}$ to four times the covariance matrix $\widehat{V}\left(\widehat{\theta}_{OLS}\right)$.

With regard to the prior specification of α_0 and h_0, we follow Primiceri (2005) and Benati and Mumtaz (2007). Let $P = AD^{1/2}$ be the Choleski factor of the time-invariant variance covariance matrix $\widehat{\Sigma}_{OLS}$ of the reduced-form innovations from the estimation of the fixed-coefficient VAR(4) where A is a lower triangular matrix with ones on the diagonal, and $D^{1/2}$ denotes a diagonal matrix whose elements are the standard deviations of the residuals. The prior for the log volatilities is as follows:

$$\ln h_0 \sim N\left(\ln \mu_0, 10 \times I_4\right) \tag{2}$$

where μ_0 is a vector that contains the diagonal elements of $D^{1/2}$ squared, and the variance-covariance matrix is arbitrarily set to ten times the identity matrix to make the prior only weakly informative. The prior for the contemporaneous interrelations is set to

$$\alpha_0 \sim N\left[\widetilde{\alpha}_0, \widetilde{V}\left(\widetilde{\alpha}_0\right)\right] \tag{3}$$

where the prior mean for α_0 is obtained by taking the inverse of A and stacking the elements below the diagonal row by row in a vector in the following way: $\widetilde{\alpha}_0 = [\widetilde{\alpha}_{0,21}, \widetilde{\alpha}_{0,31}, \widetilde{\alpha}_{0,32}, \widetilde{\alpha}_{0,41}, \widetilde{\alpha}_{0,42}, \widetilde{\alpha}_{0,43}]'$. The covariance matrix, $\widetilde{V}\left(\widetilde{\alpha}_0\right)$, is assumed to be diagonal, and each diagonal element is set to ten times the absolute value of the corresponding element in $\widetilde{\alpha}_0$. While this scaling is obviously arbitrary, it accounts for the relative magnitude of the elements in $\widetilde{\alpha}_0$ as pointed out by Benati and Mumtaz (2007).

With regard to the hyperparameters, we make the following assumptions along the lines of Benati and Mumtaz (2007). We postulate that Q follows an inverse-Wishart distribution:

$$Q \sim IW\left(\overline{Q}^{-1}, T_0\right) \tag{4}$$

where T_0 is the prior degrees of freedom which is set equal to the length of the training sample, which is sufficiently long (25 years of quarterly data) to guarantee a proper prior. Following Primiceri (2005), we adopt a relatively conservative prior for the time variation in the parameters in setting the scale matrix to $\overline{Q} = (0.01)^2 \cdot \widehat{V}\left(\widehat{\theta}_{OLS}\right)$ multiplied by the prior degrees of freedom. This is a weakly informative prior, and the particular choice for its starting value is not expected to influence the results substantially since the prior is dominated by the sample information as time progresses. We experimented with different initial conditions inducing a different amount of time variation in the coefficients to test whether our results were sensitive to the choice of the prior specification. We follow Cogley

and Sargent (2005) in setting the prior degrees of freedom alternatively to the minimum value allowed for the prior to be proper, $T_0 = \dim(\theta_t) + 1$, together with a smaller value of the scale matrix, $\overline{Q} = 3.5e^{-4} \cdot \widehat{V}\left(\widehat{\theta}_{OLS}\right)$, which together put little weight on our prior belief about the drift in θ_t. Our results are not materially affected by the different choices for this prior.

The three blocks of S are assumed to follow inverse-Wishart distributions, with the prior degrees of freedom set equal to the minimum value required for the prior to be proper:

$$S_i \sim IW\left(\overline{S}_i^{-1}, i+1\right) \tag{5}$$

where $i = 1, 2, 3$ indexes the blocks of S. The scale matrices are calibrated on the absolute values of the respective elements in $\widetilde{\alpha}_0$ as in Benati and Mumtaz (2007). Specifically, $\bar{S}_i$ is a diagonal matrix with the relevant elements of $\widetilde{\alpha}_0$ multiplied by 10^{-3}.

Given the univariate feature of the laws of motion of the stochastic volatilities, the variances of the innovations to the univariate stochastic volatility equations are drawn from an inverse-Gamma distribution as in Cogley and Sargent (2005):

$$\sigma_i^2 \sim IG\left(\frac{10^{-4}}{2}, \frac{1}{2}\right) \tag{6}$$

This distribution is proper and has fat tails.

B.2 Markov Chain Monte Carlo algorithm for simulating the posterior distribution

Since sampling from the joint posterior is complicated, we simulate the posterior distribution by sequentially drawing from the conditional posterior of the four blocks of parameters: the coefficients θ^T, the simultaneous relations A^T, the variances H^T, where the superscript T refers to the whole sample, and the hyperparameters – the elements of Q, S and the σ_i^2 for $i = 1, ..., 4$ – collectively referred to as M. Posteriors for each block of the Gibbs sampler are conditional on the observed data Y^T and the rest of the parameters drawn at previous steps.

Step 1: Drawing coefficient states

Conditional on A^T, H^T, M and Y^T, the measurement equation is linear and has Gaussian innovations with known variance. Therefore, the conditional posterior is a product of

Gaussian densities, and θ^T can be drawn using a standard simulation smoother (see Carter and Kohn 1994; Cogley and Sargent 2002) which produces a trajectory of parameters:

$$p\left(\theta^T \mid Y^T, A^T, H^T\right) = p\left(\theta_T \mid Y^T, A^T, H^T\right) \prod_{t=1}^{T-1} p\left(\theta_t \mid \theta_{t+1}, Y^T, A^T, H^T\right) \tag{7}$$

From the terminal state of the forward Kalman filter, the backward recursions produce the required smoothed draws that take the information of the whole sample into account. More specifically, the last iteration of the filter provides the conditional mean $\theta_{T|T}$ and conditional variance $P_{T|T}$ of the posterior distribution. A draw from this distribution provides the input for the backward recursion at $T-1$, $T-2$ and so on until the beginning of the sample according to:

$$\begin{aligned} \theta_{t|t+1} &= \theta_{t|t} + P_{t|t}P_{t+1|t}^{-1}\left(\theta_{t+1} - \theta_t\right) \\ P_{t|t+1} &= P_{t|t} - P_{t|t}P_{t+1|t}^{-1}P_{t|t} \end{aligned} \tag{8}$$

Step 2: Drawing covariance states

Similarly, the posterior of A^Tconditional on θ^T, H^T and Y^T is a product of normal densities and can be calculated by applying the same algorithm as in step 1 as a consquence of the block diagonal structure of the variance-covariance matrix S. More specifically, a system of unrelated regressions based on the relation $A_t u_t = \varepsilon_t$, where ε_t are orthogonalized innovations with known time-varying variance H_t and $u_t = y_t - X_t'\theta_t$ are observable residuals, can be estimated to recover A^T according to the following transformed equations where the residuals are independent standard normal:

$$\begin{aligned} u_{1,t} &= \varepsilon_{1,t} \\ \left(h_{2,t}^{-\frac{1}{2}}u_{2,t}\right) &= -\alpha_{2,1}\left(h_{2,t}^{-\frac{1}{2}}u_{1,t}\right) + \left(h_{2,t}^{-\frac{1}{2}}\varepsilon_{2,t}\right) \\ \left(h_{3,t}^{-\frac{1}{2}}u_{3,t}\right) &= -\alpha_{3,1}\left(h_{3,t}^{-\frac{1}{2}}u_{1,t}\right) - \alpha_{3,2}\left(h_{3,t}^{-\frac{1}{2}}u_{2,t}\right) + \left(h_{3,t}^{-\frac{1}{2}}\varepsilon_{3,t}\right) \\ \left(h_{4,t}^{-\frac{1}{2}}u_{4,t}\right) &= -\alpha_{4,1}\left(h_{4,t}^{-\frac{1}{2}}u_{1,t}\right) - \alpha_{4,2}\left(h_{4,t}^{-\frac{1}{2}}u_{2,t}\right) - \alpha_{4,3}\left(h_{4,t}^{-\frac{1}{2}}u_{3,t}\right) + \left(h_{4,t}^{-\frac{1}{2}}\varepsilon_{4,t}\right) \end{aligned} \tag{9}$$

Step 3: Drawing volatility states

Conditional on θ^T, A^T and Y^T, the orthogonalized innovations $\varepsilon_t \equiv A_t\left(y_t - X_t'\theta_t\right)$ with $Var\left(\varepsilon_t\right) = H_t$ are observable. However, drawing from the conditional posterior of H^T is more involved because the conditional state-space representation for $\ln h_{i,t}$ is not Gaussian. The log-normal prior on the volatility parameters is common in the stochastic volatility literature

but such a prior is not conjugate. Following Cogley and Sargent (2005, Appendix B.2.5) and Benati and Mumtaz (2007), we apply the univariate algorithm by Jacquier, Polson, and Rossi (1994) that draws the volatility states $h_{i,t}$ one at a time.

Step 4: Drawing hyperparameters

The hyperparameters M of the model can be drawn directly from their respective posterior distributions since the disturbance terms of the transition equations are observable given θ^T, A^T, H^T and Y^T.

We perform $100,000$ iterations of the Gibbs sampler and discard the first $50,000$ draws as "burn-in". The remaining sequence of draws from the conditional posteriors of the four blocks form a sample from the joint posterior distribution $p\left(\theta^T, A^T, H^T, M \mid Y^T\right)$. We keep only every 10^{th} draw in order to mitigate the autocorrelation among the draws. Following Primiceri (2005) and Benati and Mumtaz (2007), we ascertain that our Markov chain has converged to the ergodic distribution by computing the draws' inefficiency factors which are the inverse of the relative numerical efficiency measure (RNE) introduced by Geweke (1992),

$$RNE = (2\pi)^{-1} \frac{1}{S(0)} \int_{-\pi}^{\pi} S(\omega) d\omega \tag{10}$$

where $S(\omega)$ is the spectral density of the retained draws from the Gibbs sampling replications for each set of parameters at frequency ω.[4] Figure 1A displays the inefficiency factors for the states and the hyperparameters of the model, which are all far below the value of 20 designated as an upper bound by Primiceri (2005). Thus, the autocorrelation across draws is modest for all elements, which provides evidence of convergence to the ergodic distribution. In total, we collect $5,000$ simulated values from the Gibbs chain on which we base our structural analysis.

C Impulse responses and sign restrictions

Here we describe the Monte Carlo integration procedure that we use to compute the impulse response functions to a structural oil supply shock. In the spirit of Koop, Pesaran, and Potter (1996), we compute the generalized impulse responses as the difference between the conditional expectations with and without the exogenous shock:

$$IRF_{t+k} = E\left[y_{t+k} \mid \varepsilon_t, \omega_t\right] - E\left[y_{t+k} \mid \omega_t\right] \tag{11}$$

[4] See Benati and Mumtaz (2007) for details on the implementation.

where y_{t+k} contains the forecasts of the endogenous variables at horizon k; ω_t represents the current information set, and ε_t is the current disturbance term. At each point in time, the information set upon which we condition the forecasts contains the actual values of the lagged endogenous variables and a random draw of the model parameters and hyperparameters. More specifically, in order to calculate the conditional expectations we simulate the model in the following way. We randomly draw from the Gibbs sampler output one possible state of the economy at time t represented by the time-varying lagged coefficients and the elements of the variance-covariance matrix. Starting from this random draw from the joint posterior that includes the hyperparameters, we employ the transition laws and stochastically simulate the future paths of the coefficient vector and the components of the variance-covariance matrix for up to 20 quarters into the future. By projecting the evolution of the system in this way, we account for all the potential sources of uncertainty deriving from the additive innovations, variations in the lagged coefficients and changes in the contemporaneous relations among the variables in the system.

To obtain the time-varying structural impact matrix $B_{0,t}$, we implement the procedure proposed by Rubio-Ramírez, Waggoner, and Zha (2010). Given the current state of the economy, let $\Omega_t = P_t D_t P_t'$ be the eigenvalue-eigenvector decomposition of the VAR's time-varying variance covariance matrix Ω_t at time t. We draw an $N \times N$ matrix, K, from the $N(0,1)$ distribution, take the QR decomposition of K, where R is a diagonal matrix whose elements are normalized to be positive, and Q is a rotation matrix the columns of which are orthogonal to each other, and compute the time-varying structural impact matrix as $B_{0,t} = P_t D_t^{\frac{1}{2}} Q'$. Given this contemporaneous impact matrix, we compute the reduced-form innovations based on the relationship $u_t = B_{0,t}\varepsilon_t$, where ε_t contains four structural shocks drawn from a standard normal distribution. Impulse responses are then computed by comparing the effects of a shock on the evolution of the endogenous variables to the benchmark case without a shock, where in the former case the shock is set to $\varepsilon_{i,t} + 1$, while in the latter we only consider $\varepsilon_{i,t}$. The reason for this is to allow the system to be impacted by other disturbances during the propagation of the shock of interest. From the set of impulse responses derived in this way, we select only those impulse responses that at horizons $t + k$, $k = 0, 1, ..., 4$, satisfy the sign restrictions, i.e., the responses of the endogenous variables are consistent with the structural shock we wish to identify; all others are discarded.

We repeat this procedure until 100 iterations have fulfilled the sign restrictions for a given simulated future path of the economy, and then calculate the mean responses of the four endogenous variables over the accepted rotations. For each point in time, we randomly draw

500 current states of the economy that provide the distribution of impulse responses taking into account possible developments of the structure of the economy. The representative impulse response function for each variable at each date is the median of this distribution.

D A Monte Carlo study

To explore whether our econometric model with smooth transitions is well suited to capture abrupt changes in the data, we carry out a Monte Carlo exercise based on simulated data where the underlying data-generating process is characterized by a one-time break. Given that our benchmark model is too complex to be amenable to a Monte Carlo study, we assess the performance of its main building blocks by conducting two separate experiments. We generate data from (1) an AR(1) model with one exogenous regressor that features a one-off regime shift in its coefficients, and (2) a bivariate version of our benchmark VAR(4) model with an abrupt break in the variance. These two simpler models provides a parsimonious way to assess the appropriateness of modeling structural change in a smoothly evolving way as opposed to a regime switch.

D.1 A regression model with a break in the coefficients

To illustrate the effects of incorrectly assuming a smooth process for the evolution of the coefficients, we simulate data from the following stationary AR(1) model with one exogenous regressor, written in demeaned form:

$$y_t = \alpha_i y_{t-1} + \beta_i x_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0,1) \tag{12}$$

where we set $\alpha_1 = 0.2$ and $\beta_1 = 0.5$ for the first half of the sample and $\alpha_2 = 0.6$ and $\beta_2 = 1.5$ for the second half of the sample. For each sample generated with this parameterization, we estimate a model that postulates that the coefficient vector $\delta = [\alpha \;\; \beta]'$ evolves smoothly according to a driftless random walk process:

$$\delta_t = \delta_{t-1} + \eta_t \qquad \eta_t \sim N(0, Q) \tag{13}$$

We estimate the state-space model in equations (12) and (13) by Bayesian methods described in Kim and Nelson (1999). The unrestricted prior for the initial state is Gaussian:

$$\delta_0 \sim N(\widehat{\delta}_{OLS}, 4 \cdot \widehat{V}(\widehat{\delta}_{OLS})) \tag{14}$$

where $\widehat{\delta}_{OLS}$ and $\widehat{V}(\widehat{\delta}_{OLS})$ are the OLS point estimate and asymptotic variance based on a training sample as in our benchmark model. For the variance σ^2 in the observation equation, we postulate an inverse-gamma distribution:

$$\sigma^2 \sim IG\left(\frac{\lambda}{2}, \frac{\nu}{2}\right) \tag{15}$$

with scale parameter $\lambda = 0.01$ and degrees-of-freedom parameter $\nu = 2$. The prior for Q is assumed to be inverse Wishart:

$$Q \sim IW\left(\overline{Q}^{-1}, \xi\right) \tag{16}$$

where $\overline{Q} = 0.01 \cdot \xi$ and $\xi = 3$. The starting values for the coefficients are set to the OLS estimates, $\sigma_0 = 1$, and $Q_0 = 0.1 \cdot I_2$ where I_2 is a 2×2 identity matrix. The time-varying coefficients are drawn using the Carter and Kohn (1994) algorithm outlined above. We constrain α_t to be less than one in absolute value at all dates t. The first $2,000$ draws in the Gibbs simulation process are discarded to ensure convergence. The posterior mean of $\widehat{\delta}_t$ is computed based on the remaining $1,000$ generated values. To evaluate how well this model can pick up the break imposed in the data-generating process, we also obtain an estimate of $\widehat{\delta}_t^{dummy}$ from a model that includes a dummy variable that takes a value of 0 before the break and 1 thereafter. In this way, we can construct error bands that capture the parameter uncertainty in estimating the true model.

We construct sample sizes of $T_1 = 200$ and $T_2 = 600$ after discarding the first $1,000$ periods to remove the influence of initial values. A sample size of 200 can be considered the equivalent of the typical sample length for quarterly time series available for the post-WWII period and 600 is representative of such a dataset at monthly frequency. There are $T/2$ data points on each side of the break date. We carry out $1,000$ Monte Carlo replications for each model and sample size.

Figure 2A reports the mean of the estimates for the exogenous coefficient and for the AR coefficient for the smooth-transition model and for the discrete-break model together with the 68% and 90% posterior credible sets for the two sample sizes. The estimation results show that the drifiting coefficient model locates the break in a satisfactory manner and moves relatively swiftly to the new regime.

D.2 A bivariate VAR model with a break in the variance

In the second experiment, the data-generating process is a bivariate VAR(4) model similar to equation (1) in the main text:

$$y_t = X_t'\theta + \varepsilon_t \tag{17}$$

where y_t denotes a vector of variables; X_t is a matrix including four lags of y_t and a constant; θ is a coefficient matrix, and $\varepsilon_t \sim N(0, \Omega_i)$, $i = 1, 2$ with the following variance-covariance matrices for two subperiods:

$$\Omega_1 = \begin{bmatrix} 20 & 5 \\ 5 & 30 \end{bmatrix} \qquad \Omega_2 = \begin{bmatrix} 1.5 & -4 \\ -4 & 300 \end{bmatrix} \tag{18}$$

To obtain a realistic parameterization for Ω_1 and Ω_2, we take guidance from the estimation of a bivariate VAR model for oil production and the real price of oil over two subsamples. The data generated from this model mimic a specific feature of the observed oil production and oil price series, namely a considerable decrease in oil production volatility and an increase in oil price volatility after the break in the variance. Figure 3A, panel A illustrates this behavior in one such random sample. The length of each sample generated from this model is 750, and the initial 500 periods are removed to yield a sample similar in size to that used in the empirical analysis.

For each sample, we estimate the time-varying VAR model with stochastic volatility presented in section B. We retain the same priors as in the benchmark model and obtain initial values from the estimation of a constant-coefficient VAR(4) over a twenty–five-year training sample. This leaves us with 150 observations for the actual estimation, and the regime switch occurs at $t = 68$. Given the greater complexity of this model, we can only perform 250 Monte Carlo replications, and the results should consequently be viewed as suggestive. It should, however, be sufficient to examine the speed of transition from one regime to the other which is the main feature of interest.

Figure 3A, panel B displays the time profile of the average of the variance estimates over the Monte Carlo simulations together with both the 16^{th} and 84^{th} and the 5^{th} and 95^{th} percentiles of the posterior distribution. The results indicate that our approach has the power to detect the regime shift to a satisfactory degree even in a relatively short sample.

E Further analysis

E.1 Evidence for time variation

In assessing the relative importance of time variation over the sample, we consider the joint posterior distribution of impulse responses across selected pairs of oil market episodes presented in a scatterplot along the lines of Cogley, Primiceri, and Sargent (2010). Shifts of this distribution away from the 45-degree line are indicative of a systematic change over time. Figure 4A reports the joint posterior distributions of the impact response of oil production and the cumulated responses of real GDP and CPI four quarters after an oil supply shock normalized such that it raises the real price of oil by 10% on impact for pairs of representative dates. Figure 5A reports the joint posterior distributions for the case when the oil supply shock corresponds to a 1% shortfall in oil production on impact for the same combinations of dates. Values for the earlier date are always plotted on the x-axis and those for the later date on the y-axis so that the location of the joint distribution of positive (negative) responses above the 45-degree line indicates an increase (decrease) over time and below the 45-degree line a decrease (increase) over time.

The evidence for time variation is most compelling for world oil production. The joint posterior draws for almost all combinations of dates are clustered above the dividing line, pointing towards a systematic decrease in the responses to oil supply shocks as time progresses. The joint posterior distributions for the real price of oil indicate a systematic increase in the magnitude of the impact of an oil supply shock. The exception is the pair $1979Q3 : 1986Q1$ for which the pairwise draws are more scattered suggesting no significant difference.

For real GDP, the points of the joint distribution are almost equally spread out around the 45-degree line for pairs of dates that are not too far apart from each other for both normalizations. This suggests the absence of a significant change in the reaction of real GDP to oil supply shocks. There is, however, some evidence for differences between the magnitudes of output responses during episodes that are more distant in time. In particular, an oil supply shock normalized on oil production is more contractionary in the more recent past compared to early periods since a considerable fraction of the pairwise draws lies below the threshold. The opposite results emerges for the normalization based on the real price of oil. While the dispersion of pairwise posterior draws for consumer prices implies similar responses for pairs of dates in the early part of the sample, there appears to be a systematic

increase in price responses in the last decade with the majority of draws lying above the diving line for both normalizations. Overall, this provides strong evidence for sizeable changes in the responses of oil market variables and some support for changes in the responses of the aggregate economy over time.

E.2 Sensitivity analysis

We summarize the results for the alternative model specifications discussed in section 3.3 in the main text and conduct two additional robustness checks. In what follows, an oil supply shock is normalized to correspond to a 1% decrease in world oil production.

Figure 6A, panel A displays the time-varying median responses of the unemployment rate (left) and the implicit GDP deflator (right) which yield the same pattern of time variation as real GDP and CPI. For ease of comparison, the dotted lines in panels B and C and in Figure 7A depict the median estimates obtained with the baseline model. Figure 6A, panel B, shows that the time-varying responses obtained with different oil price measures are essentially identical. Panel C presents the evolution of the responses for the model augmented by the federal funds rate which exhibit a pattern remarkably similar to the baseline case. Figure 7A, panel A displays the time-varying responses when sign restrictions are imposed from $t = 4$ to $t = 8$. As before, the time-varying responses closely track the dotted lines that represent the benchmark model, demonstrating that our main conclusions are not sensitive to the modified identification assumptions.

As pointed out by Fry and Pagan (2011), sign restrictions impose only weak information. Building on this insight, Kilian and Murphy (2011) argue that the sign identification strategy needs to be complemented with additional information in order to derive economically meaningful results in the context of oil market models. They propose the use of empirically plausible boundary restrictions on the magnitudes of the implied short-run price elasticities of oil demand and oil supply as auxiliary identification criteria to eliminate those structural models that are associated with implausibly high elasticities. To verify the robustness of our findings, we follow Kilian and Murphy (2010) in imposing that the short-run price elasticity for oil demand cannot exceed its long-run counterpart which may be inferred to be about -0.8 using cross-sectional evidence from U.S. household surveys (see e.g. Hausman and Newey 1995). Figure 7A, panel B shows that our baseline results are not affected by this additional identifying restriction. Even more stringent bounds on the impact price elasticity of oil demand have little effect on the time profile of the impulse responses.

To explore the sensitivity of our results with regard to the data frequency, we estimate a monthly VAR(12) model that includes the growth rates of world oil production, the real refiners' acquisition cost of crude oil imports, U.S. industrial production, and U.S. CPI over two subsamples, $1974M1 - 1985M12$ and $1986M1 - 2011M3$.[5] In line with the quarterly model, the sign restrictions are postulated to hold over a horizon of 12 months after the shock. Figure 8A displays the median responses of the four variables to an oil supply shock together with the 16^{th} and 84^{th} percentiles of the posterior distribution. The results for the monthly specification paint much the same picture of time variation as in the quarterly split sample model.

References

[1] **Benati, Luca, and Haroon Mumtaz.** 2007. "U.S. Evolving Macroeconomic Dynamics: A Structural Investigation." ECB Working Paper 746.

[2] **Carter, Christopher K., and Robert Kohn.** 1994. "On Gibbs Sampling for State Space Models." *Biometrika*, 81(3): 541-53.

[3] **Cogley, Timothy, Giorgio E. Primiceri, and Thomas J. Sargent.** 2010. "Inflation-Gap Persistence in the US." *American Economic Journal: Macroeconomics*, 2(1): 43-69.

[4] **Cogley, Timothy, and Thomas J. Sargent.** 2002. "Evolving Post-World War II U.S. Inflation Dynamics." In *NBER Macroeconomics Annual 2001*, ed. Ben S. Bernanke and Kenneth Rogoff, 331-73. Cambridge, MA: MIT Press.

[5] **Cogley, Timothy, and Thomas J. Sargent.** 2005. "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US." *Review of Economic Dynamics*, 8(2): 262-302.

[6] **Energy Information Administration**. 2010. *Annual Energy Review.*

[7] **Energy Information Administration**. 2011 (August). *Monthly Energy Review.*

[5] Estimating the benchmark time-varying VAR model with monthly data is not an option since this would result in a proliferation of free parameters given the number of lags that are required to allow for sufficient dynamics.

[8] **Energy Information Administration**. 2011 (June). *Short-Term Energy Outlook*.

[9] **Fry, Renée, and Adrian Pagan.** 2011. "Sign Restrictions in Structural Vector Autoregressions: A Critical Review." *Journal of Economic Literature*, forthcoming.

[10] **Geweke, John.** 1992. "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In: *Bayesian Statistics*, ed. José M. Bernardo, James O. Berger, Alexander P. Dawid and A.F.M. Smith, 169-93. Oxford University Press, Oxford.

[11] **Hausman, Jerry A., and Whitney K. Newey**. 1995. "Nonparametric Estimation of Exact Consumers' Surplus and Deadweight Loss." *Econometrica*, 63: 1445-76.

[12] **IMF**. 2006 (August). *World Economic Outlook*.

[13] **Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi.** 1994. "Bayesian Analysis of Stochastic Volatility Models." *Journal of Business and Economic Statistics*, 12(4): 371-418.

[14] **Kilian, Lutz, and Daniel P. Murphy**. 2011. "Why Agnostic Sign Restrictions Are Not Enough: Understanding the Dynamics of Oil Market VAR Models." *Journal of the European Economic Association*, forthcoming.

[15] **Kilian, Lutz, and Daniel P. Murphy**. 2010. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil." Mimeo, University of Michigan.

[16] **Kim, Chang-Jin, and Charles R. Nelson.** 1999. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge, MA: MIT Press.

[17] **Koop, Gary, M. Hashem Pesaran, and Simon M. Potter.** 1996. "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics*, 74(1): 119-47.

[18] **Litterman, Robert B.** 1983. "A Random Walk, Markov Model for the Distribution of Time Series." *Journal of Business and Economic Statistics*, 1(2): 169-73.

[19] **Mork, Knut Anton.** 1989. "Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results." *Journal of Political Economy*, 97(3): 740-44.

[20] *Oil & Gas Journal*. Weekly, various issues for 1953-1972. Tulsa, Oklahoma: Pennwell Corporation.

[21] **Primiceri, Giorgio E.** 2005. "Time Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economic Studies*, 72(3): 821-52.

[22] **Rubio-Ramírez, Juan, Dan F. Waggoner, and Tao Zha.** 2010. "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference." *Review of Economic Studies*, 77(2): 665-96.

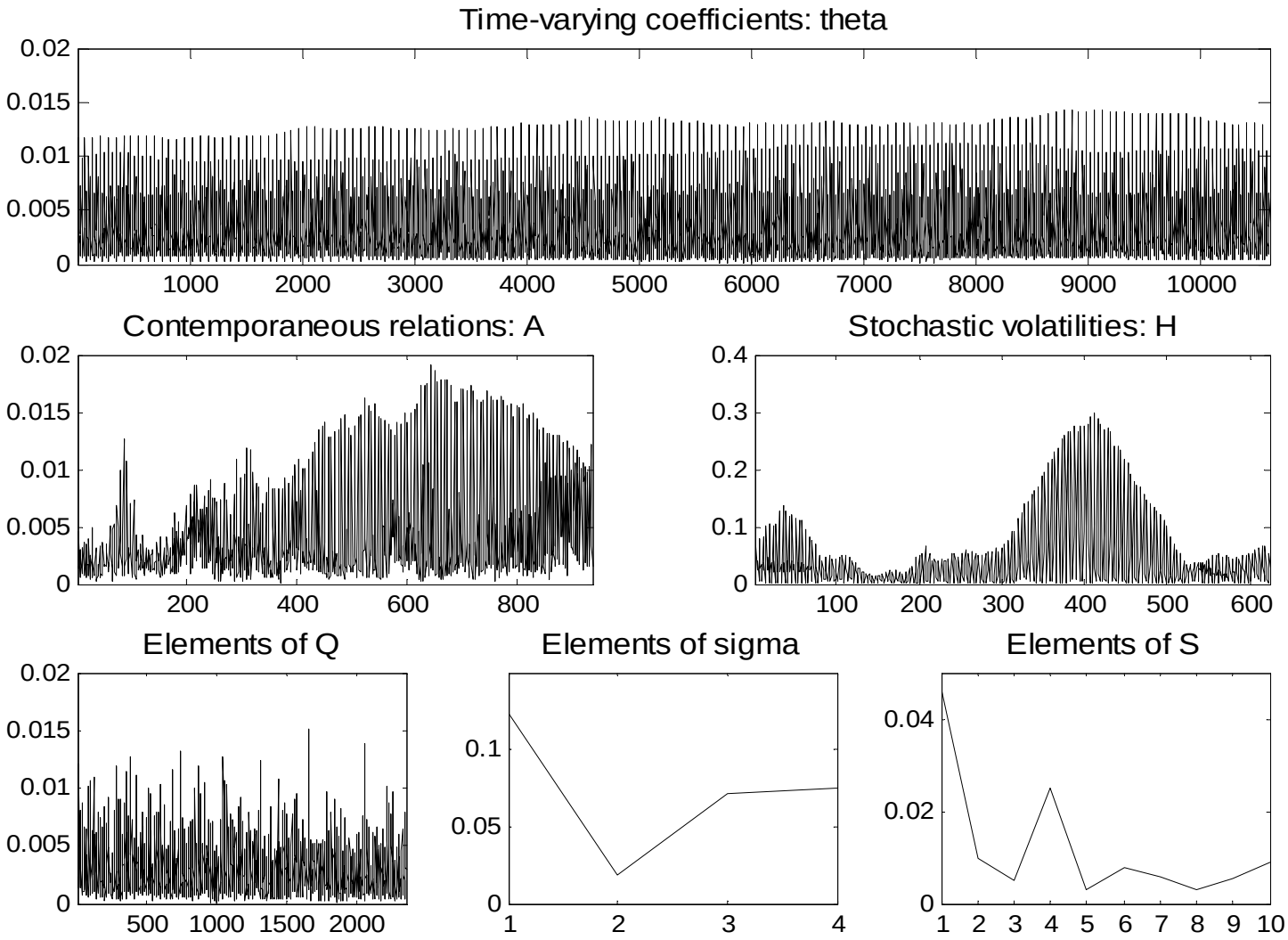
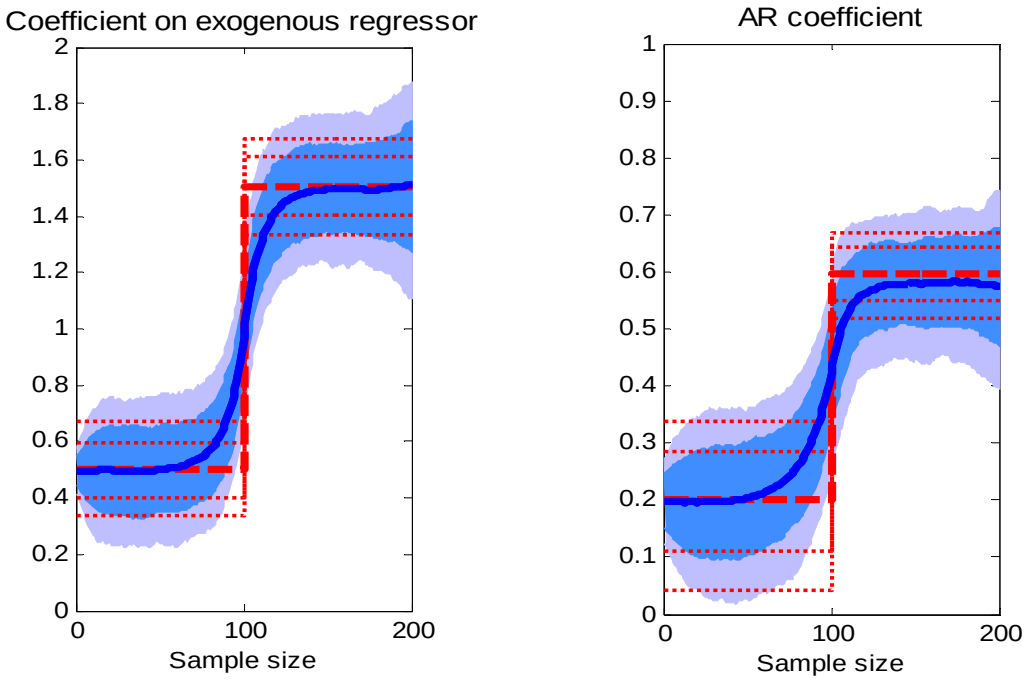


Figure 1A: Assessing the convergence of the Markov chain: inefficiency factors for the draws from the ergodic distribution for the states and hyperparameters.

Panel A



Panel B

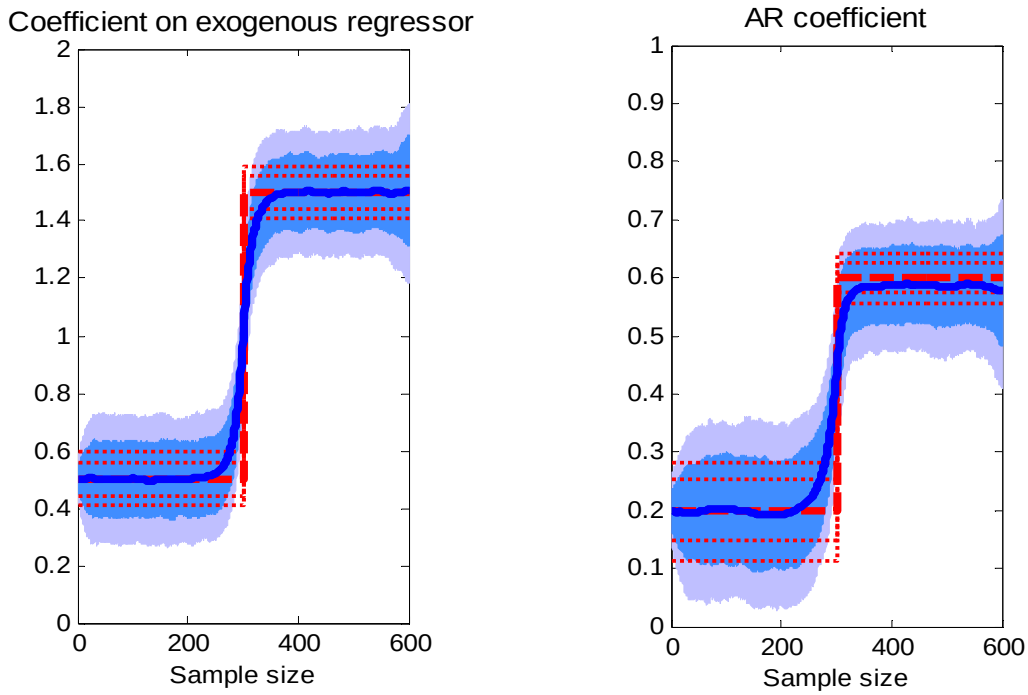
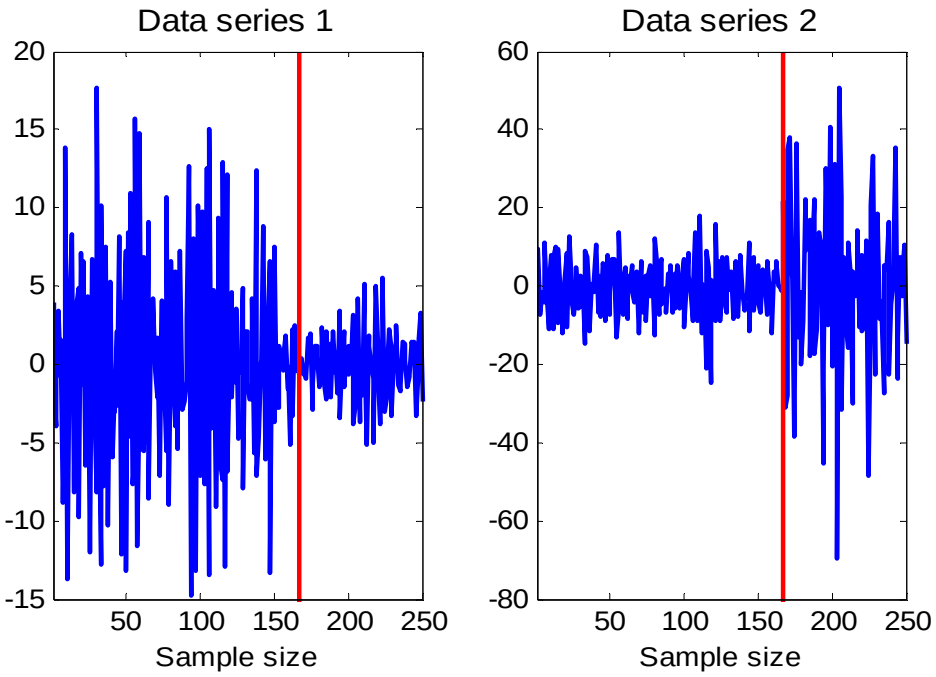


Figure 2A: Mean, 68% and 90% posterior credible sets of coefficient estimates from a smooth-transition model (solid line and shaded areas) and a discrete-break model (dashed and dotted lines) for 1,000 Monte Carlo replications.
Panel A: Sample size *T=200*.
Panel B: Sample size *T=600*.

Panel A



Panel B

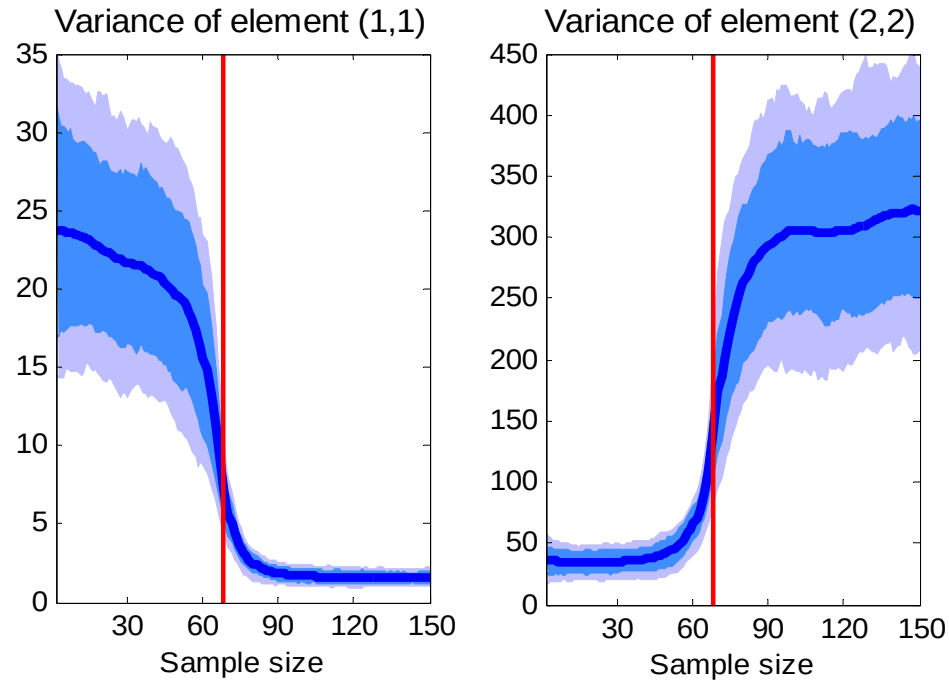


Figure 3A: Panel A: Random sample generated from bivariate VAR(4) model with break in variance at *t=167* (vertical line).
Panel B: Mean, 68% and 90% posterior credible sets of variance estimates for 250 Monte Carlo replications.

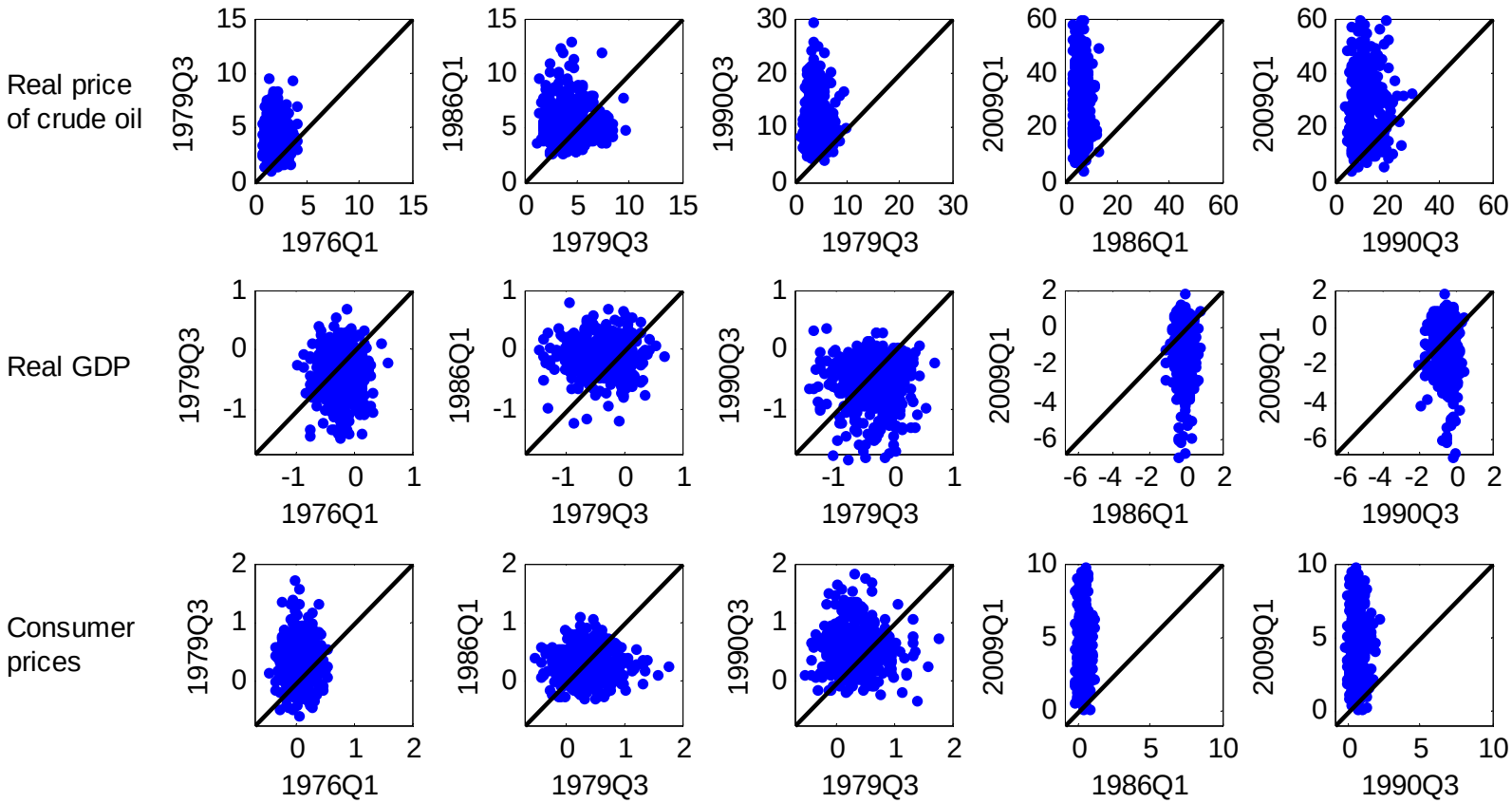


Figure 4A: Joint posterior distribution of the responses of the real price of oil, real GDP and consumer prices to an oil supply shock normalized such that it decreases world oil production by 1% on impact for selected pairs of dates.

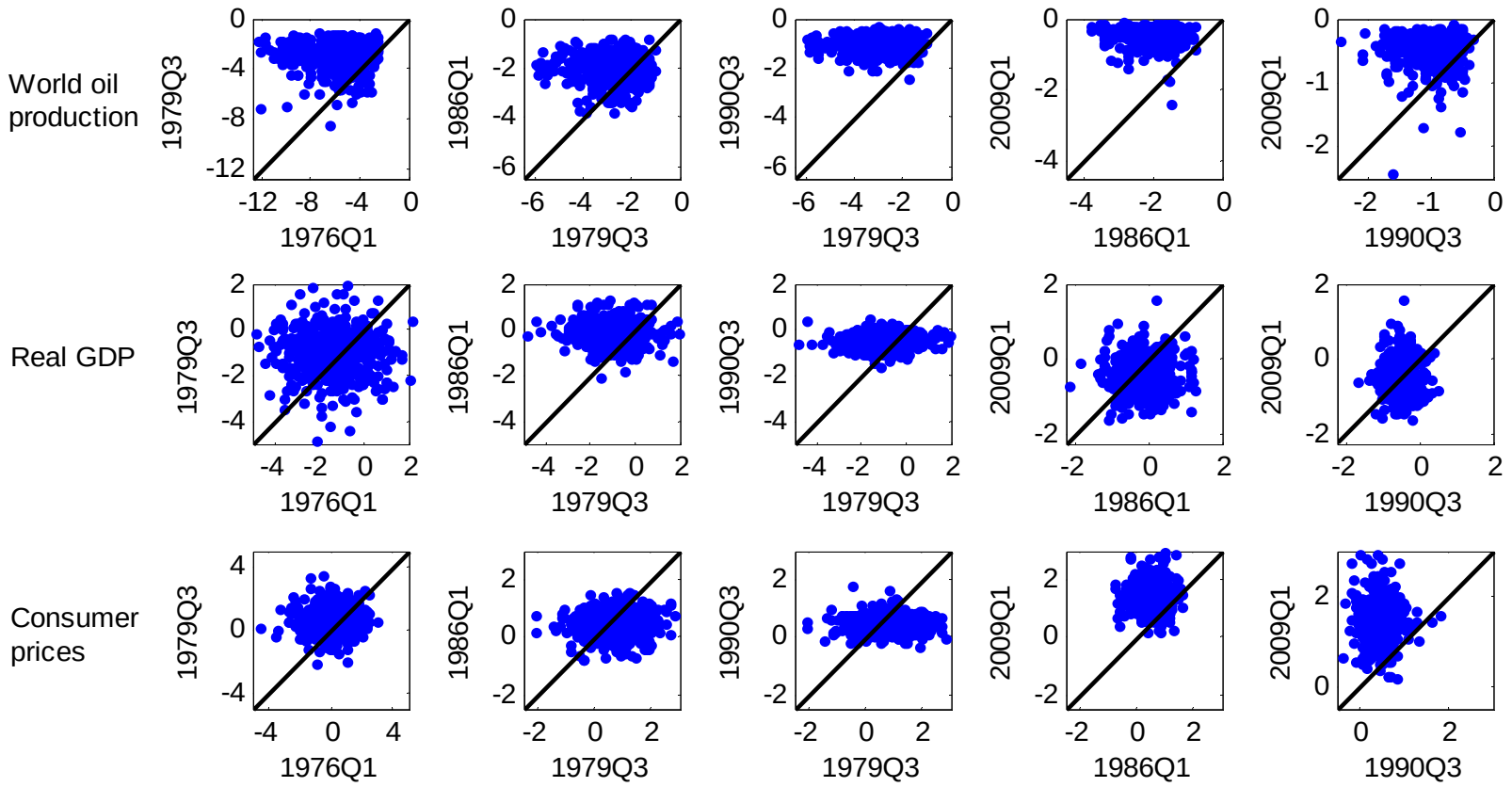


Figure 5A: Joint posterior distribution of the responses of world oil production, real GDP and consumer prices to an oil supply shock normalized such that it raises the real price of oil by 10% on impact for selected pairs of dates.

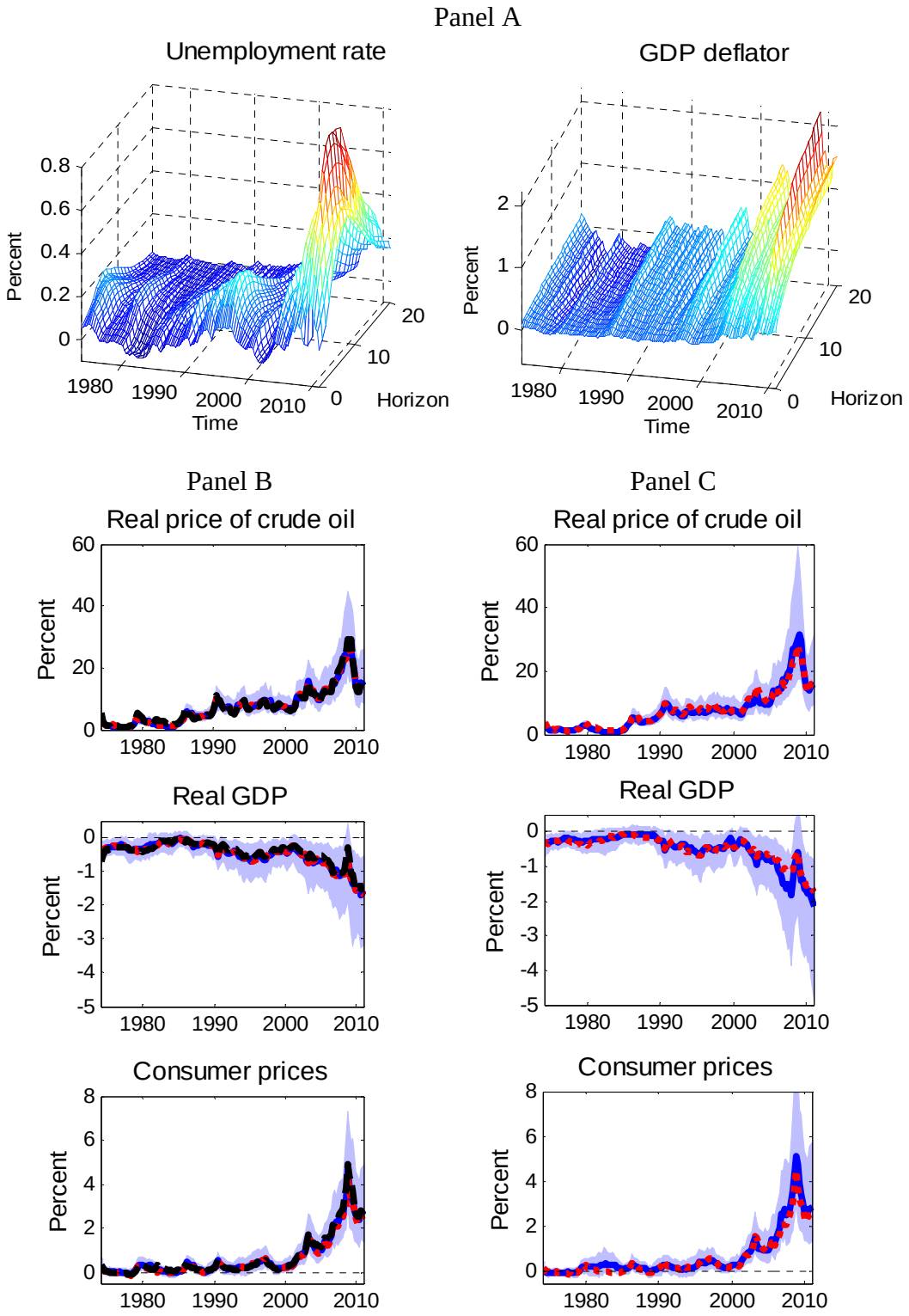


Figure 6A: Panel A: Median responses of U.S. unemployment (left) and GDP deflator (right).
Panel B: Median responses for model with WTI (dashed line) and composite refiners' acquisition cost (solid line) with 68% posterior credible set (shaded area).
Panel C: Median responses for specification with federal funds rate.

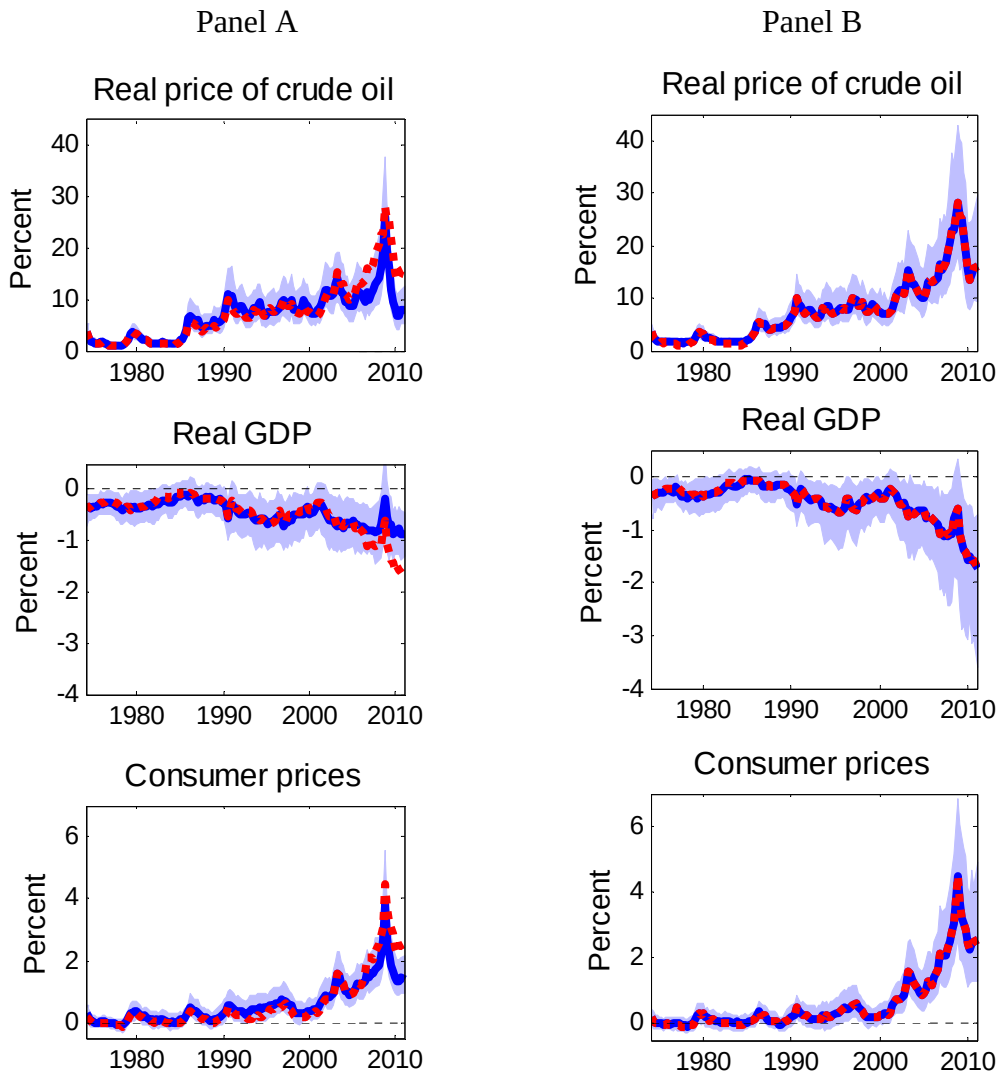


Figure 7A: Panel A: Median responses with 16th and 84th percentiles when sign restrictions are imposed from *t=4* to *t=8*.
Panel B: Median responses with 16th and 84th percentiles when the lower bound for the short-run price elasticity of demand is -0.8

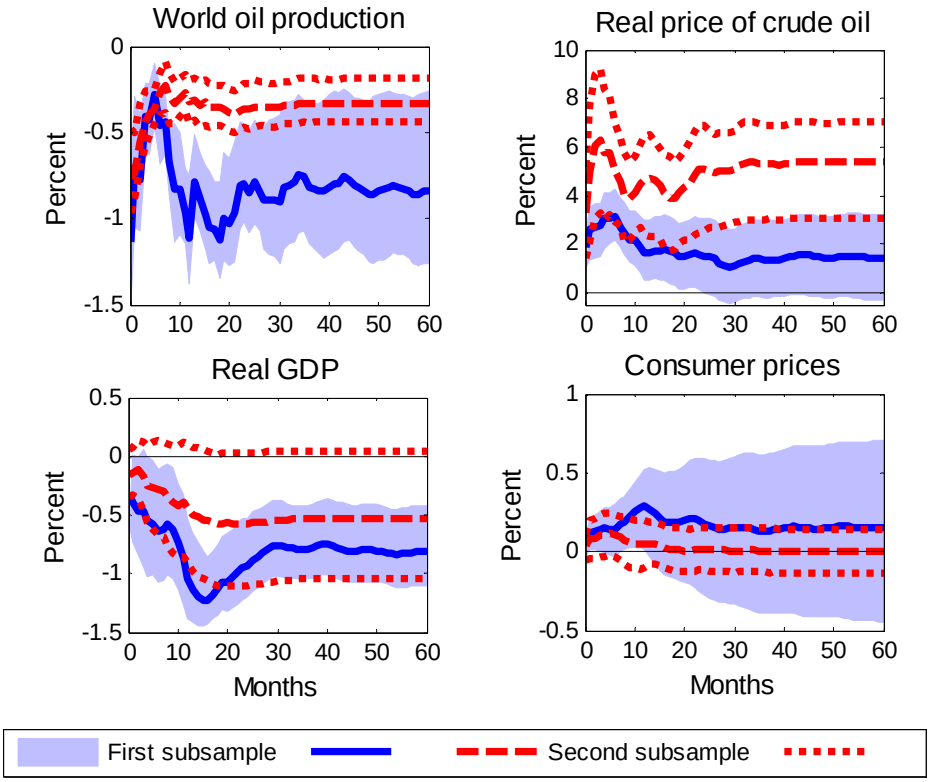


Figure 8A: Median responses with 16th and 84th percentiles obtained with a constant-coefficient VAR estimated over two subsamples 1974*M1*–1985*M12* and 1986*M1*–2011*M3*.