Technical Appendix: Policy Uncertainty and Aggregate Fluctuations.

Haroon Mumtaz Paolo Surico

July 18, 2017

1 The Gibbs sampling algorithm

Prior Distributions and starting values

Consider the model to be estimated

$$Z_t = c + \sum_{j=1}^{P} \beta_j Z_{t-j} + \sum_{j=0}^{J} \gamma_j \tilde{h}_{t-j} + \Omega_t^{1/2} e_t, e_t \tilde{} N(0,1) \quad (1)$$

$$\Omega_t = A^{-1} H_t A^{-1'}, H_t = diag\left(\exp \tilde{h}_t\right) \quad (2)$$

$$\tilde{h}_t = \theta \tilde{h}_{t-1} + Q^{1/2} \eta_t \quad , \eta_t \tilde{} N(0,1), \quad E\left(e_t, \eta_{i,t}\right) = 0, \quad i = 1, 2..N \quad (3)$$

VAR coefficients

Let the vectorised coefficients of equation 1 be denoted by $\Gamma = vec\left(\beta_j, \gamma_j, c\right)$. The initial conditions for the VAR coefficients Γ_0 (to be used in the Kalman filter as described below) are obtained via an OLS estimate of equation (1) using an initial estimate of the stochastic volatility. The covariance around these initial conditions P_0 is set to a diagonal matrix with diagonal elements equal to 10.

The initial estimate of stochastic volatility is obtained via a simpler version of the benchmark model where the stochastic volatility does not enter the mean equations. We use a training sample of 40 observations to initialize the estimation of this simpler model. The Gibbs algorithm for this model is a simplified version of the algorithm described in Cogley and Sargent (2005), employing uninformative priors. The estimated volatility from this model is added as exogenous regressors to a VAR using the data described in the text in order to provide a rough guess for initial conditions for the VAR coefficients.

Elements of H_t

The prior for $\tilde{h}_t$ at $t = 0$ is defined as $\tilde{h}_0 \sim N(\ln \mu_0, I_N)$ where μ_0 are the first elements of the initial estimate of the stochastic volatility described above.

Elements of A

The prior for the off-diagonal elements A is $A_0 \sim N\left(\hat{a}, V\left(\hat{a}\right)\right)$ where $\hat{a}$ are the elements of this matrix from the initial estimation described above. $V\left(\hat{a}\right)$ is assumed to be diagonal with the elements set equal to the absolute value of the corresponding element of $\hat{a}$.

Parameters of the transition equation

We postulate a Normal, inverse-Wishart prior distribution for the coefficients and the covariance matrix of the transition equation (3). Under the prior mean, each stochastic volatility follows an AR(1) process with an AR(1) coefficient equal to the estimated value over the training sample. The prior is implemented via dummy observations (see Banbura et al (2010)) and the prior tightness is set to 0.1.

Simulating the Posterior Distributions

The joint posterior distribution $H(\Gamma, A, H_t, \theta, Q)$ is approximated via a Metropolis within Gibbs algorithm that samples from the following conditional posterior distributions:

VAR coefficients : $H(\Gamma|A, H_t, \theta, Q)$

The distribution of the VAR coefficients Γ conditional on all other parameters Ξ and the stochastic volatility $\tilde{h}_t$ is linear and Gaussian: $\Gamma|Z_t, \tilde{h}_t, \Xi \sim N\left(\Gamma_{T|T}, P_{T|T}\right)$ where $\Gamma_{T|T} = E\left(\Gamma_T|Z_t, \tilde{h}_t, \Xi\right), P_{T|T} = Cov\left(\Gamma_T|Z_t, \tilde{h}_t, \Xi\right)$. Following Carter and Kohn (1994), we use the Kalman filter to estimate $\Gamma_{T|T}$ and $P_{T|T}$ where we account for the fact that the covariance matrix of the VAR residuals changes through time. The final iteration of the Kalman filter at time T delivers $\Gamma_{T|T}$ and $P_{T|T}$. The Kalman filter is initialized using the initial conditions (Γ_0, P_0) described above. This application of Carter and Kohn's algorithm to our heteroskedastic VAR model is equivalent to a GLS transformation of the model.

Element of A : $H(A|\Gamma, H_t, \theta, Q)$

Given a draw for Γ and $\tilde{h}_t$, the VAR model can be written as $A(v_t) = e_t$ where $v_t = Z_t - c + \sum_{j=1}^{P} \beta_j Z_{t-j} + \sum_{j=0}^{J} \gamma_j \tilde{h}_{t-j}$ and $VAR(e_t) = H_t$. For a triangular A matrix, this is a system of linear equations with known form of heteroskedasticity. The conditional distributions for a linear regression apply to this system after a simple GLS transformation to make the errors homoskedastic (see Cogley and Sargent (2005)). The ith equation of this system is given as $v_{it} = -\alpha v_{-it} + e_{it}$ where the subscript i denotes the ith column while $-i$ denotes columns 1 to $i-1$. Note that the variance of e_{it} is time-varying and given by $\exp\left(\tilde{h}_{it}\right)$. A GLS transformation involves dividing both sides of the equation by $\sqrt{\exp\left(\tilde{h}_{it}\right)}$ to produce $v^*_{it} = -\alpha v^*_{-it} + e^*_{it}$ where * denotes the transformed variables and $var(e^*_{it}) = 1$. The conditional posterior for α is normal with mean and variance given by M^* and V^* :

$$\begin{aligned} M^* &= \left(V\left(\hat{a}^{ols}\right)^{-1} + v^{*\prime}_{-it} v^*_{-it}\right)^{-1} \left(V\left(\hat{a}^{ols}\right)^{-1} \hat{a}^{ols} + v^{*\prime}_{-it} v^*_{it}\right) \\ V^* &= \left(V\left(\hat{a}^{ols}\right)^{-1} + v^{*\prime}_{-it} v^*_{-it}\right)^{-1} \end{aligned}$$

The identification scheme in Blanchard and Perotti (2002) involves a non-triangular A matrix and can be written as $Cv_t = Fe_t$. However, as shown in Pereira and Lopes (2014), the C and the F matrices can be transformed such

that each implied equation only contains exogenous shocks on the right hand side. Given this transformation, Cogley and Sargent's equation by equation algorithm becomes applicable again.

Elements of H_t : $H\left(H_t|A,\Gamma,\theta,Q\right)$

Conditional on the VAR coefficients and the parameters of the transition equation, the model has a multivariate non-linear state-space representation. Carlin, Polson and Stoffer (1992) show that the conditional distribution of the state variables in a general state space model can be written as the product of three terms:

$$\tilde{h}_t|Z_t,\Xi \propto f\left(\tilde{h}_t|\tilde{h}_{t-1}\right)\times f\left(\tilde{h}_{t+1}|\tilde{h}_t\right)\times f\left(Z_t|\tilde{h}_t,\Xi\right) \tag{4}$$

where Ξ denotes all other parameters. In the context of stochastic volatility models, Jacquier, Polson and Rossi (1994) show that this density is a product of log normal densities for $\bar{h}_t$ and $\bar{h}_{t+1}$ and a normal density for Z_t where $\bar{h}_t = \exp\left(\tilde{h}_t\right)$. Carlin, Polson and Stoffer (1992) derive the general form of the mean and variance of the underlying normal density for $f\left(\tilde{h}_t|\tilde{h}_{t-1},\tilde{h}_{t+1},\Xi\right)\propto f\left(\tilde{h}_t|\tilde{h}_{t-1}\right)\times f\left(\tilde{h}_{t+1}|\tilde{h}_t\right)$ and show that this is given by:

$$f\left(\tilde{h}_t|\tilde{h}_{t-1},\tilde{h}_{t+1},\Xi\right)\tilde{}N\left(B_{2t}b_{2t},B_{2t}\right) \tag{5}$$

where $B_{2t}^{-1}=\tilde{Q}^{-1}+\tilde{F}'\tilde{Q}^{-1}\tilde{F}$ and $b_{2t}=\tilde{h}_{t-1}\tilde{F}'\tilde{Q}^{-1}+\tilde{h}_{t+1}\tilde{Q}^{-1}\tilde{F}$. Here $\tilde{F}$ and $\tilde{Q}$ denote the coefficients and the error variance of the transition equation, i.e. θ and Q in companion form. Note that, due to the non-linearity of the observation equation of the model, an analytical expression for the complete conditional $\tilde{h}_t|Z_t,\Xi$ is unavailable and a Metropolis step is required.

Following Jacquier, Polson and Rossi (1994), we draw from (4) using a date by date independence Metropolis step with the density in (5) being the candidate generating density. This choice implies that the acceptance probability is given by the ratio of the conditional likelihood $f\left(Z_t|\tilde{h}_t,\Xi\right)$ at the old and the new draw. In order to take endpoints into account, the algorithm is modified slightly for the initial condition and the last observation. Details of these changes can be found in Jacquier, Polson and Rossi (1994).

Parameters of the transition equation : $H\left(\theta|\Gamma,A,H_t,Q\right)$ and $H\left(Q|\Gamma,A,H_t,\theta\right)$

Conditional on a draw for $\tilde{h}_t$, the transition equation (3) is a VAR(1) model with a diagonal covariance matrix. The conditional posterior for the coefficients θ is normal with mean and variance given respectively by:

$$\begin{aligned}\theta^* &= \left(x^{*\prime}x^*\right)^{-1}\left(x^{*\prime}y^*\right)\\ v^* &= Q\otimes\left(x^{*\prime}x^*\right)^{-1}\end{aligned}$$

where $y^*=[\tilde{h}_t;y_d]$ and $x^*=[\tilde{h}_{t-1};x_d]$ with y_d and x_d denoting the dummy observations that implement the prior.

The conditional posterior for Q is inverse Wishart and is given by

$$H\left(Q|\Gamma,A,H_t,\theta\right)\tilde{}IW(S^*,T^*)$$

where T^* denote the number of actual observations plus the number of dummy observations and $S^*=\left(y^*-x^*b^*\right)'\left(y^*-x^*b^*\right)$

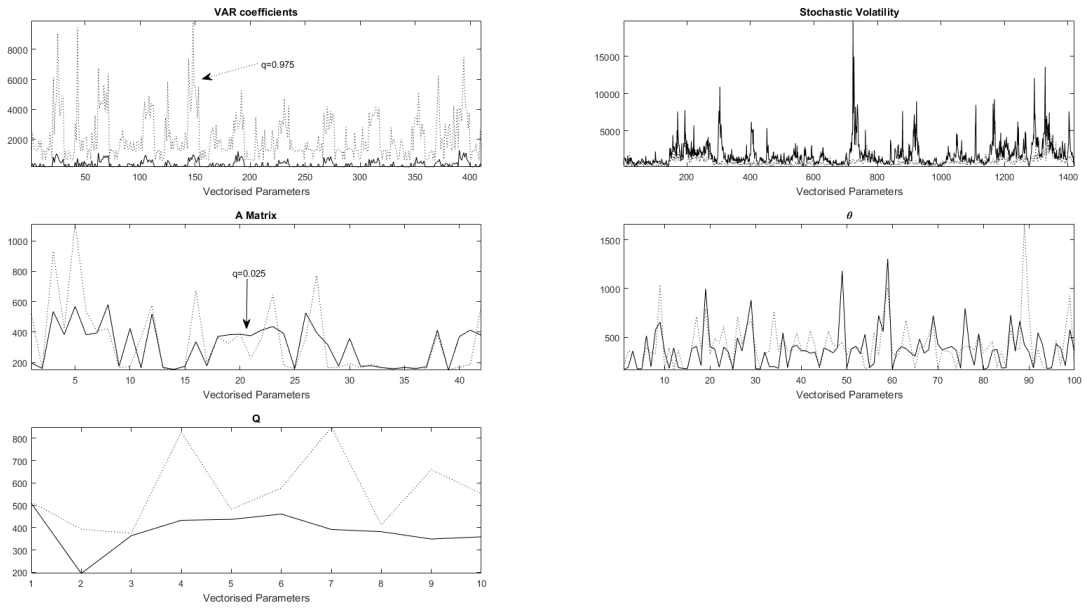

Figure 1: The Raftery and Lewis (1992) diagnostic.

The on-line technical appendix to the paper presents a small Monte-Carlo experiment that shows that this algorithm displays a satisfactory performance.

Convergence

The MCMC algorithm is applied using 500,000 iterations discarding the first 50,000 as burn-in. We retain every 45th draw out of the remaining 450,000 iterations. In order to assess convergence, we compute the Raftery and Lewis (1992) diagnostic which indicates the total length of the run required to generate a desired level of accuracy. We report the diagnostic for two quantiles 0.025 and 0.975. As in Primiceri (2005), the remaining parameters are: desired accuracy 0.025, probability of attaining desired accuracy 0.95. The results are presented in figure 1. The figure shows the estimated total length of the run across the elements of the different parameter block. Note that the suggested number of iterations are well below the 500,000 iterations employed in our algorithm. As a further check we calculate inefficiency factors (IF) and report them in figure 2. The IF are an estimate of $1 + 2\sum_{k=1}^{\infty} \rho_k$ where ρ_k is the autocorrelation of the chain and the infinite lag is approximated using a Parzen window. Values of IF around 20 are deemed acceptable. With the exception of some stochastic volatilities, this conditions seems to be satisfied for most parameters. For the stochastic volatilities the majority (greater than 70%) of IF are below 30. Given the large number of endogenous and state variables, in our view this is reasonable evidence for convergence.

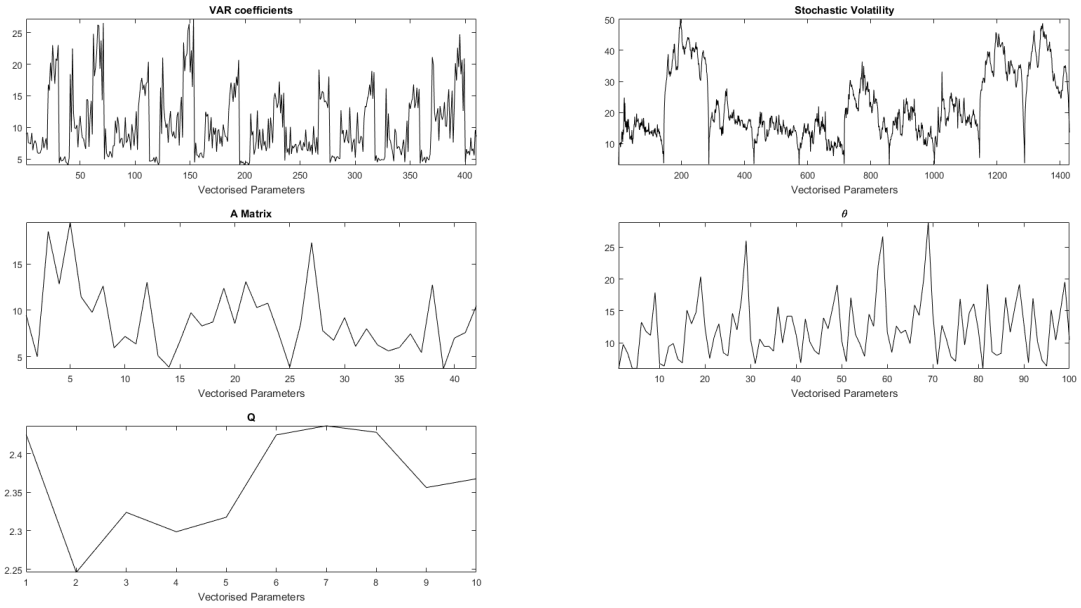

Figure 2: Inefficiency Factors

Monte-Carlo Experiment

We conduct a Monte-Carlo experiment to test the robustness of the model. The DGP is defined by the following Bi-variate VAR with stochastic volatility in mean

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} 0.5 & 0.2 \\ -0.2 & 0.5 \end{pmatrix}\begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}\begin{pmatrix} \ln h_{1t} \\ \ln h_{2t} \end{pmatrix} + \begin{pmatrix} d_1 & d_2 \\ d_3 & d_4 \end{pmatrix}\begin{pmatrix} \ln h_{1t-1} \\ \ln h_{2t-1} \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}$$

where $\begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} \tilde{} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega_t\right)$

$$\Omega_t = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} h_{1t} & 0 \\ 0 & h_{2t} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}^{-1'}$$

and

$$\begin{pmatrix} \ln h_{1t} \\ \ln h_{2t} \end{pmatrix} = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.9 \end{pmatrix}\begin{pmatrix} \ln h_{1t-1} \\ \ln h_{2t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}, \quad \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \tilde{} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, I_2\right)$$

We generate data from two versions of this DGP. In the benchmark case $\begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} -0.1 & -0.1 \\ 0.1 & 0.1 \end{pmatrix}$ and $\begin{pmatrix} d_1 & d_2 \\ d_3 & d_4 \end{pmatrix} = \begin{pmatrix} -0.2 & -0.2 \\ 0.2 & 0.2 \end{pmatrix}$. In the alternative DGP we assume $\begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} d_1 & d_2 \\ d_3 & d_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and volatility shocks have no direct impact on the endogenous variables in this case. We generate 500 observations for each DGP and discard the first 100 to remove the impact of starting values. Using this artificial data $Z_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix}$, the following model is estimated using 5000 MCMC replications (with a burn-in of 4000 iterations)

$$Z_t = c + \beta_1 Z_{t-1} + \sum_{j=0}^{1} \gamma_j \tilde{h}_{t-j} + \Omega_t^{1/2} e_t$$

where $\Omega_t = A^{-1} H_t A^{-1'}$ and $\tilde{h}_t = [\ln h_{1t}, \ln h_{2t}], H_t = diag\left(\exp\left(\tilde{h}_t\right)\right)$. The two stochastic volatilities are assumed to evolve as AR(1) processes

$$\tilde{h}_t = \theta \tilde{h}_{t-1} + Q^{1/2} \eta_t$$

where θ, Q are assumed to be diagonal. We use identical flat priors in each case. At each replication, we estimate the linear approximation to the impulse response of Z_t to 1 standard deviation shocks to η_t. The experiment is repeated using 100 replications. Figure 3 presents the results when the benchmark DGP is used. The figure shows that the estimated response to uncertainty shocks is close to the true estimate. Figure 4 presents these estimates using the alternative DGP where uncertainty shocks do not have a direct impact on the endogenous variables. It is clear from 4 that the estimates indicate that the uncertainty shocks have no systematic impact on the endogenous variables. These results provide strong evidence that the model is able to adequately capture the impact of uncertainty shocks and that the proposed MCMC algorithm displays a reasonable performance.

2 Financial Market uncertainty

In order to account for financial market uncertainty, we re-estimate our benchmark model replacing consumer confidence with the stock market index. The volatility of the shock to this variable then represents a proxy for financial market uncertainty that is allowed to affect the endogenous variables along with policy uncertainty. Figure 5 presents the impulse responses to policy uncertainty shocks from this model and shows that the key results are preserved. Debt uncertainty is still important in this model.

3 DIC calculation

We compare the fit of the benchmark SVAR with stochastic volatility with a linear SVAR using the deviance information criterion (DIC) (see Spiegelhalter et.al. 2002). The DIC rewards model fit and penalises model complexity. The DIC is defined as:

$$DIC = \bar{D} + p_D. \tag{6}$$

The first term is $\bar{D} = E\left(-2\ln f\left(\tilde{y}\backslash\Xi^m\right)\right) \approx \frac{1}{M}\sum_m\left(-2\ln f\left(\tilde{y}\backslash\Xi^m\right)\right)$ where $f\left(\tilde{y}\backslash\Xi^m\right)$ is the likelihood evaluated at the draws from the Gibbs sampler Ξ^m where $m = 1, 2, ...M$ indexes the draws. This term measures goodness of fit. The second term p_D is defined as a measure of the number of effective parameters in the model (or model complexity). This is defined as $p_D = \bar{D} - D\left(\bar{\Xi}\right) = E\left(-2\ln f\left(\tilde{y}\backslash\Xi^m\right)\right) - \left(-2\ln f\left(\tilde{y}\backslash E(\Xi^m)\right)\right)$ and can be approximated as $p_D = \frac{1}{M}\sum_m\left(-2\ln f\left(\tilde{y}\backslash\Xi^m\right)\right) - \left(-2\ln L\left(\frac{1}{M}\sum_m \Xi^m\right)\right)$. The model with the lowest DIC is preferred. In order to evaluate the likelihood for the benchmark VAR with stochastic volatility we use a particle filter. The estimated DIC for the benchmark model is 1983.3 with an estimated goodness of fit $\bar{D} = 1290.45$ and model complexity $p_D = 692.8$. The model complexity of linear VAR, in contrast, is much lower $p_D = 264.9$. However, the fit in this model drops dramatically with $\bar{D}$ rising to 1737.78 and thus more than off-setting the reduction in model complexity. The estimated DIC for the linear VAR is 2002.75 which is larger than that for the benchmark VAR with stochastic volatility. The benchmark model is thus preferred.

4 Response to level shocks

We show the impulse responses of key variables to 1 unit level shocks in figure 6. The red line and the grey shaded area present the response from the benchmark VAR with stochastic volatility. The dotted line is the response from a simple linear VAR. In both cases, the benchmark identification scheme is used. The response to spending and tax shocks from the benchmark model resemble those obtained in the recent literature. For example, in a highly cited survey paper Caldara and Kamps (2008) finds that a spending shock raises GDP and consumption, while the impact on investment and inflation is close to zero. We reach very similar conclusions. Similarly, Caldara and Kamps (2008) reports that a tax shock has a small negative impact on GDP, a result that resembles our estimates. While our estimated response to the monetary policy shock is imprecise (perhaps due to the large number of parameters in the

model), the median responses resemble those from standard recursive VARs. Real variables decline in response to monetary tightening at about the one year horizon, but there is a price puzzle.

A comparison of the benchmark responses with those obtained using the simple VAR model suggests a similarity in the direction of the response at short horizons, but differences in propagation, with the linear responses more persistent than the non-linear in many cases. For example, GDP falls in response to tax shocks in both the benchmark and linear case. However, the median response is more persistent in the linear model. This is perhaps because the linear model does not account for the direct impact of stochastic volatility and uncertainty shocks and perhaps suffers from omitted variables as a result.

5 The response of the level of debt to policy uncertainty shocks

Figure 7 presents the response of the level of debt to policy uncertainty shocks.

References

[1] Caldara, Dario & Kamps, Christophe, 2008."What are the effects of fiscal policy shocks? A VAR-based comparative analysis," Working Paper Series 0877, European Central Bank.

[2] Spiegelhalter, David J.; Best, Nicola G.; Carlin, Bradley P.; van der Linde, Angelika (2002). "Bayesian measures of model complexity and fit (with discussion)". Journal of the Royal Statistical Society, Series B. 64 (4): 583–639. doi:10.1111/1467-9868.00353. JSTOR 3088806. MR 1979380.

6 Figures and Tables

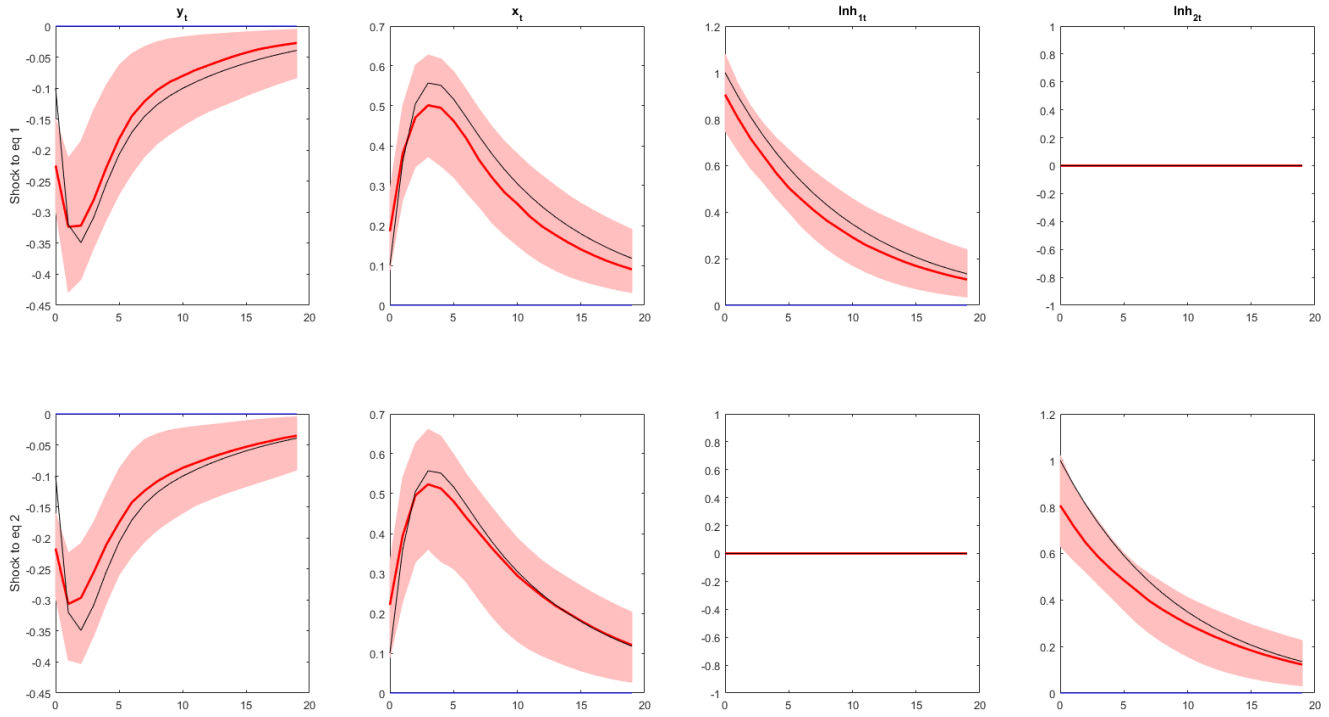

Figure 3: Monte-Carlo results using the benchmark DGP. The red line is the median across Monte-Carlo replications. The shaded area is the 1 SD band across Monte-Carlo replications. The true value of the IRFs is depicted by the black line.

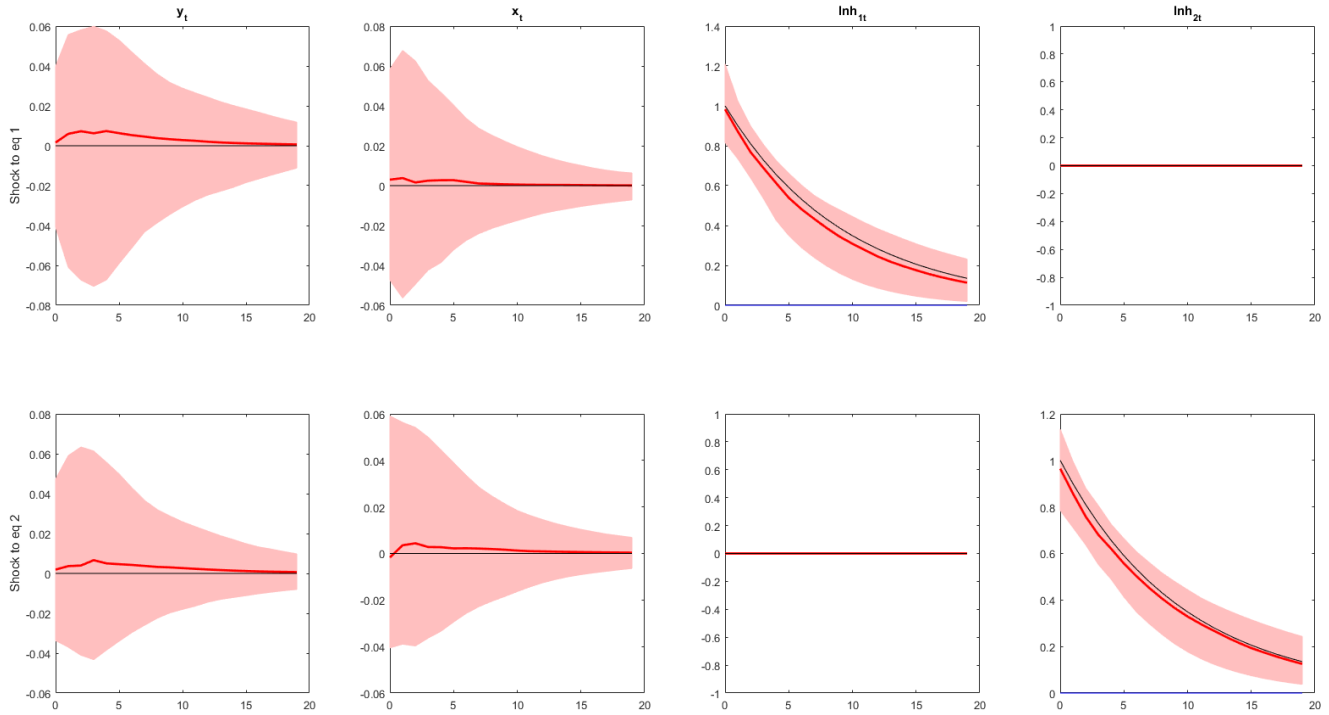

Figure 4: Monte-Carlo results using the alternative DGP. The red line is the median across Monte-Carlo replications. The shaded area is the 1 SD band across Monte-Carlo replications. The true value of the IRFs is depicted by the black line.

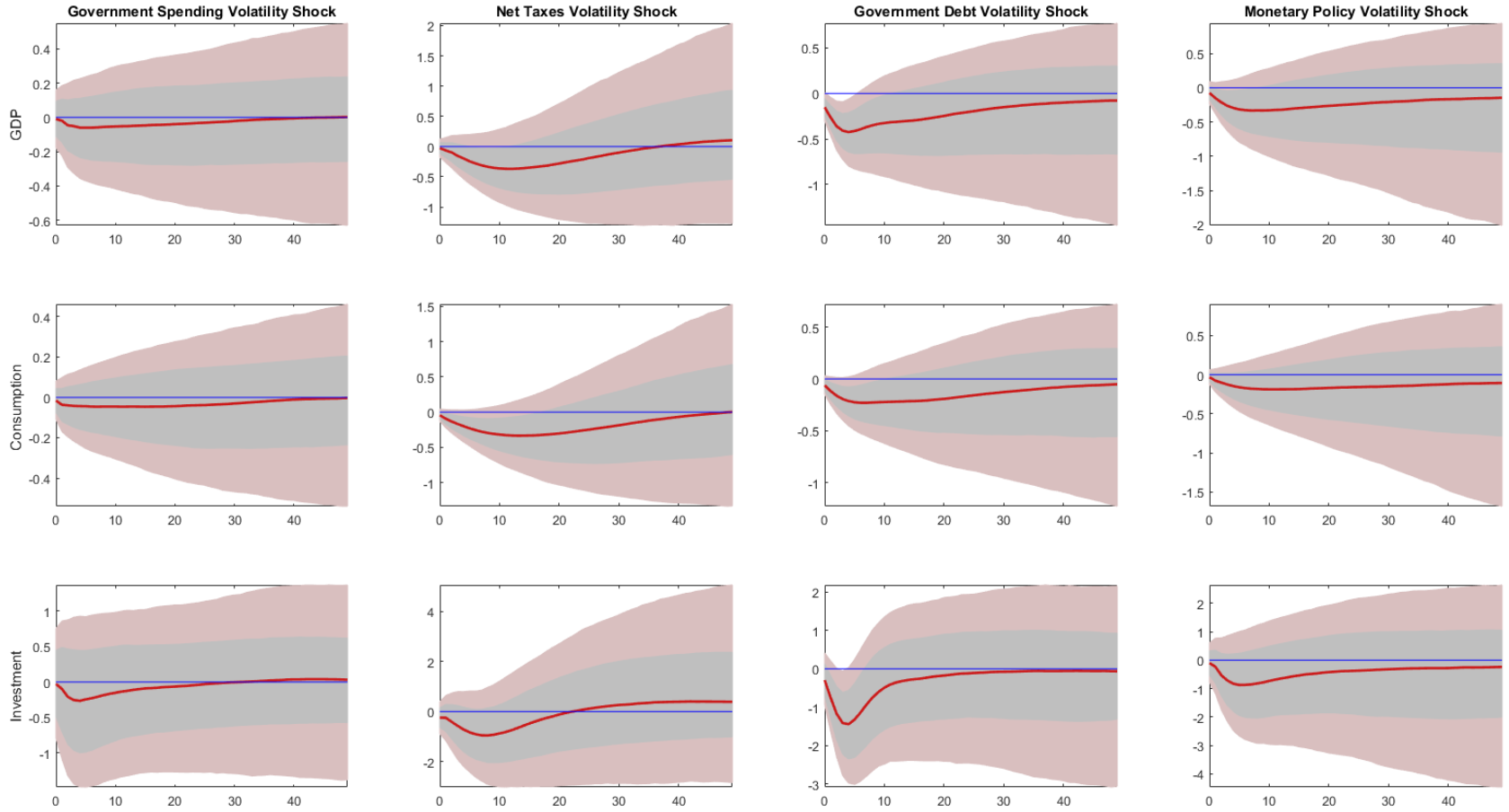

Figure 5: Impulse response of key variables to uncertainty shocks. Including the stock price index

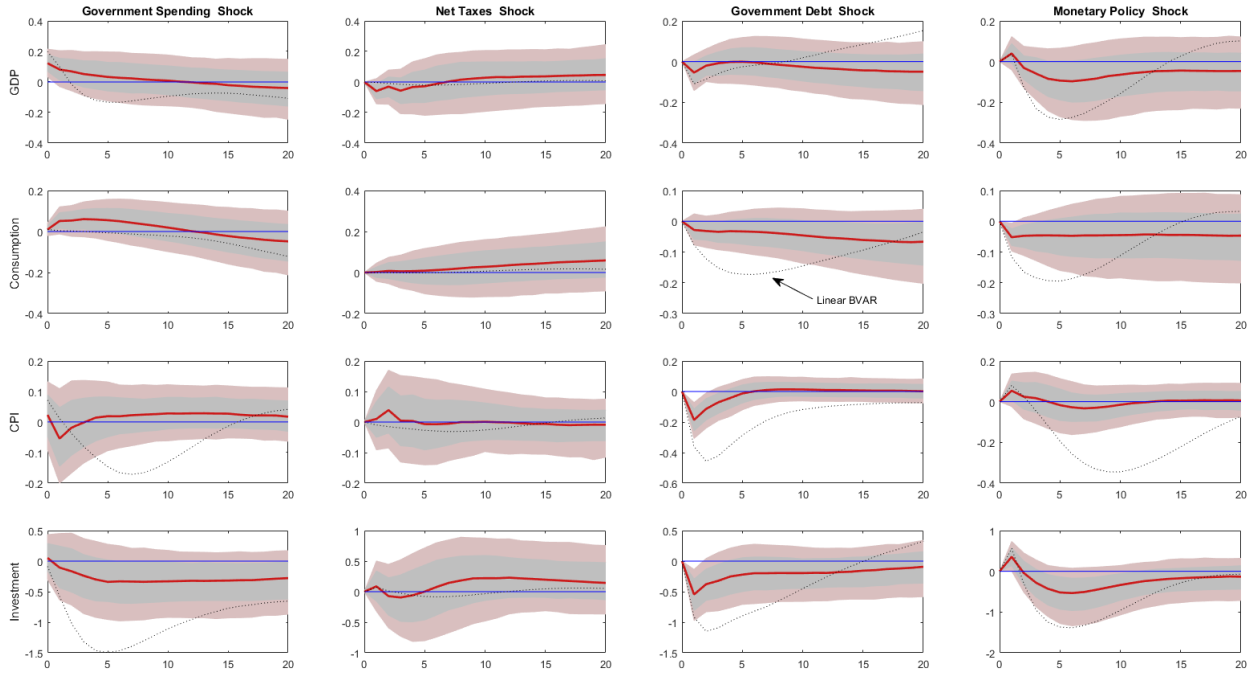

Figure 6: Impulse response to level shocks. The red line is the median. The dark grey shaded area is the 68% error band while the light shaded area is the 90% error band. The dotted black line is the response from a linear BVAR using the same identification scheme as the non-linear model.

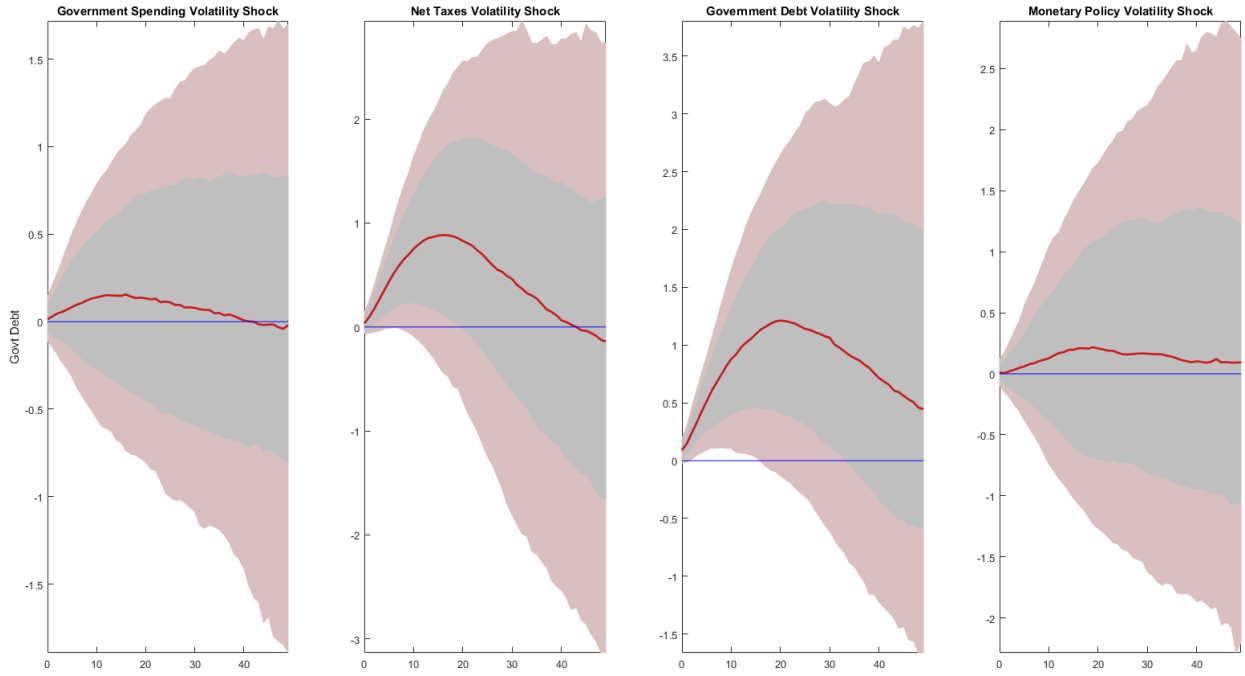

Figure 7: The response of debt to policy uncertainty shocks. The red line is the median. The dark grey shaded area is the 68% error band while the light shaded area is the 90% error band