Volatility Co-movement and the Great Moderation. An Empirical Analysis

Haroon Mumtaz and Konstantinos Theodoridis

Working Paper No. 804 October 2016 ISSN 1473-0278

School of Economics and Finance

Queen Mary University of London

Volatility co-movement and the Great Moderation. An empirical analysis*

Haroon Mumtaz† Queen Mary University

Konstantinos Theodoridis Bank of England

October 2016

Abstract

We propose an extended time-varying parameter Vector Autoregression that allows for an evolving relationship between the variances of the shocks. Using this model, we show that the relationship between the conditional variance of GDP growth and the long-term interest rate has become weaker over time in the US. Similarly, the co-movement between the variance of the long-term interest rate across the US and the UK declined over the 'Great Moderation' period. In contrast, the volatility of US and UK GDP growth appears to have become increasingly correlated in the recent past.

JEL Classification: C15, C32, E32.

Keywords: Vector-Autoregressions, Time-Varying parameters, Stochastic Volatility.

1 Introduction

A voluminous literature has shown that the volatility of US macroeconomic variables declined during the early and the mid-1980s. Prominent papers that reached this conclusion include Kim and Nelson (1999), Cogley and Sargent (2005). The former paper employs a time-vaying VAR with stochastic volatility (TVP-SVOL) to show that the conditional variance of unemployment declined by about 40 percent after the early 1980s. Their estimates also suggest a fall in inflation and short-term interest rate conditional volatility around that time. However, it is unclear if this 'great moderation' extended to a larger set of variables and that the relationship amongst conditional variances of different variables was of a similar sign and magnitude. For example, Campbell (2005) provides evidence that the volatility of stock returns was largely unaffected by this structural change and remained high after the mid-1980s. Fuentes-Albero (2012) shows that the great moderation did not extend to corporate balance sheet variables and Mumtaz and Sunder-Plassmann (2013) provide evidence that the conditional volatility of the real exchange rate increased after the 1980s while macroeconomic volatility was declining.

The recent DSGE literature has also provided economic intuition for the possibility of cross-country correlation in conditional volatilities. For example Mumtaz and Theodoridis (2015) show that when agents in a two country model have Epstein-Zin preferences, the conditional volatility of endogenous variables is time-varying and transfer of resources across countries in the face of a

*Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the Bank of England or to state Bank of England policy. This paper should therefore not be reported as representing the views of the Bank of England or members of the Monetary Policy Committee, Financial Policy Committee or Prudential Regulation Authority Board.

†Email:h.mumtaz@qmul.ac.uk

shock can induce co-movement in the conditional variances. Mumtaz and Theodoridis (2015) argue that the strength of this co-movement can vary over time and is affected by factors such as trade openness, the practice of monetary policy and the degree of price stickiness.

The aim of the current paper is to develop an empirical model that can be used to investigate the dynamic relationship between conditional volatilities. We attempt to fulfill this aim by extending the work-horse TVP-SVOL model. In particular, we focus on the transition equation for the stochastic volatilties. In contrast to the previous literature, we assume that this transition equation is itself a VAR with time-varying parameters and heteroscedasticity. This allows the computation of measures of time-varying correlation between the conditional variances at frequencies of interest. We provide a Gibbs sampling algorithm for this extended model.

We use this model to study the relationship between the conditional volatility of output growth and the long-term government bond yield in the US and the UK. Our key results suggest that the correlation between the conditional volatility of output across countries rose substantially after the mid-1980s and is estimated to be close to 1 over the recent past. In contrast, the correlation between the conditional variance of the long term interest rate has declined across the two countries. There is also evidence that within each country the conditional variance of output growth and the long term rate has become less correlated over the great moderation period. We argue that these estimates are consistent with structural changes such as increasing trade openness and a higher weight placed on inflation control by central bankers.

This paper adds to the literature on the great moderation by explicitly considering the evolving dynamic relationship between conditional volatilities. In contrast, earlier papers have largely focussed on investigating changes in the relationship amongst the levels of macroeconomic and financial variables (see for e.g. Prieto *et al.* (2016)). By considering the co-movement in second moments we shed new light on the consequences of policies and/or structural changes that resulted in the great moderation. Given that the recent financial crisis has resulted in high volatility, these results are of immediate relevance to policy makers.

From an econometric point of view, the paper makes a contribution by proposing a multivariate stochastic volatility model where the law of motion of the volatilities is characterised by time-varying parameters. This adds to the literature on stochastic volatility models that allow for jumps in the level of volatility via a Markov switching specification (see Mike K. P. So (1998) and Lopes and Carvalho (2007) for e.g.). The key advantage of our specification is that it also allows the dynamic relationship between the volatilties to evolve over time. This feature also distinguishes our contribution from alternatives such as multi-variate GARCH models, where the relationship amongst the conditional variances is constant. Similarly, recently proposed factor models with time-varying volatility (see Mumtaz and Theodoridis (2015) and Berger *et al.* (2014)) are able to capture volatility co-movement but do not directly incorporate the possibility of a time-varying dynamic relationship between the conditional variances.

The paper is organised as follows. The next section introduces the empirical model and outlines the estimation algorithm. Section 3 discusses the data used in the study, the model specification and the key empirical results. Section 4 concludes. Technical details on model estimation and selection are presented in Appendices A and B.

2 Empirical model

We estimate the following time-varying VAR model with stochastic volatility

$$Z_t = c_t + \sum_{j=1}^{P} \beta_{tj} Z_{t-j} + \Omega_t^{1/2} e_t, e_t \sim N(0,1) \tag{1}$$

where Z_t is a matrix of N endogenous variables, c_t is a time-varying vector of N intercepts and β_{tj} denotes $N \times N$ coefficient matrices at each point in time. Denoting the vectorised coefficients by $\Gamma_t = vec([\beta_t, c_t])$, the law of motion for these states is defined as

$$\Gamma_t = \Gamma_{t-1} + \tau_t, \tau_t \tilde{} N(0, Q_\Gamma) \tag{2}$$

Following Primiceri (2005), the time-varying covariance matrix Ω_t is factored as

$$\Omega_t = A_t^{-1} H_t A_t^{-1'} \tag{3}$$

where A is a $N \times N$ lower triangular matrix while $H_t = diag\left(\exp\left(\tilde{h}_t\right)\right)$ with $\tilde{h}_t = [h_{1t}, h_{2t}, ..h_{N,t}]$ denoting the stochastic volatility of each of the N orthogonal shocks in the VAR model. As in Primiceri (2005) the free elements of A_t (denoted by a_t) evolve as random walks

$$a_t = a_{t-1} + \bar{s}_t, VAR(\bar{s}_t) = \breve{S} \tag{4}$$

where $\breve{S}$ is assumed to be block diagonal.

The specification for the stochastic volatilities marks our point of departure from previous studies such as Cogley and Sargent (2005), Primiceri (2005) and Canova and Perez Forero (2015). In contrast to previous applications of the TVP-SVOL, the law of motion of the stochastic volatilties is assumed to be a VAR(1) model with time-varying parameters:

$$\tilde{h}_t = \alpha_t + \theta_t \tilde{h}_{t-1} + \eta_t, \eta_t \sim N(0, Q_t) \tag{5}$$

In equation 5, α_t denotes the N time-varying intercepts, while θ_t is a $N \times N$ matrix of time-varying slope coefficients. Letting $\Phi_t = vec([\theta_t, \alpha_t])$, the transition equation for the coefficients is given by

$$\Phi_t = \Phi_{t-1} + u_t, u_t \tilde{} N(0, Q_\Phi) \tag{6}$$

Following Primiceri (2005), we assume that the covariance matrix Q_t is non-diagonal. However, in our application this covariance matrix is time-varying. This time-variation is introduced in the covariance matrix Q_t by factoring it as $Q_t = C_t^{-1} D_t C_t^{-1'}$. C_t is a lower triangular matrix where the non-zero and non-one elements c_t evolve as random walks

$$c_t = c_{t-1} + s_t \tag{7}$$

where $var(s_t) = S$ is a block-diagonal covariance matrix. Similarly, the log of non-zero elements of the diagonal matrix D_t follow a random walk

$$d_t = d_{t-1} + n_t, n_t \sim N(0, q) \tag{8}$$

where $d_t = [d_{1t}, d_{2t}, ..d_{N,t}]$ and $D_t = diag(\exp(d_t))$.

These extensions to the basic TVP-SVOL model are in a similar spirit to the modifications to

this model introduced in Cogley *et al.* (2010). These authors introduce time-variation in the law of motion for the VAR coefficients in order to allow for the possibility of a shift in the dynamics of Γ_t. In the present application our interest centers on volatility co-movement and a consequence we extend the law of motion for $\tilde{h}_t$ to incorporate co-movement and time-variation. Given an estimate of Φ_t and Q_t, it is trivial to compute an approximation to the unconditional correlation matrix of $\tilde{h}_t$ at each point in time. In addition, the multivariate spectrum of $\tilde{h}_t$ can also be approximated and the relationship amongst the volatilities at different frequencies can be examined.

2.1 Estimation

The model is estimated via a Gibbs sampling algorithm. The details of the algorithm are provided in Appendix A with a brief description of the main steps given below. The appendix also provides details of all the prior distributions. Note that a key feature of the latter is the prior density for Q_Φ. Following the TVP-VAR literature this prior is set using a training sample. In particular, we estimate a VAR with stochastic volatility that excludes time-varying dynamics in the transition equation to get a rough estimate for $\tilde{h}_t$. Then using a training sample of T_0 observations we estimate a fixed coefficient VAR(1) model using this initial estimate of $\tilde{h}_t$ as endogenous variables and obtain the OLS estimate of the coefficient covariance matrix $\hat{Q}_{OLS}$. The prior for Q_Φ is then set as an inverse Wishart density with scale parameter $\hat{Q}_{OLS} \times T_0 \times \kappa$ and degrees of freedom $\dim(\Phi) + n$. Following Cogley and Sargent (2005), the scaling parameter κ is set equal to $3.5e-04$. We set $n = 10$ implying some weight on the prior. We find that while using a more non-informative prior (i.e. with $n = 1$) results in very similar estimates of the time-varying correlations, the changes in some of the parameters are estimated to be very volatile.[1]We use non-informative priors for the blocks of S. The scale matrix of this inverse Wishart prior is set to a matrix with diagonal elements given by 10^{-3} and the degrees of freedom are the dimension of the block plus 1.

The Gibbs algorithm samples from the following conditional posterior distributions.

1. $G\left(\Gamma_t \backslash A_t, \tilde{h}_t, Q_\Gamma\right)$. Given $\tilde{h}_t$, the model in equation 1 is a linear Gaussian state-space model. We draw from this conditional posterior using the Carter and Kohn (1994) algorithm.

2. $G\left(A_t \backslash \Gamma_t, \tilde{h}_t, \breve{S}\right)$. Given the VAR coefficients, the model can be written in terms of the residuals as $A_t u_t = H_t^{1/2} e_t$. This defines a series of heteroscedastic, time-varying regressions in the residuals. Given $\tilde{h}_t$ this again represents a series of linear, Gaussian state-space models and the Carter and Kohn (1994) algorithm can again be employed.

3. $G(\Phi_t \backslash \tilde{h}_t, C_t, D_t, Q_\Phi)$. Given $\tilde{h}_t$ and Q_t, equations 5 and 6 constitute a linear Gaussian state-space model. We draw from the conditional posterior of Φ_t using the Carter and Kohn (1994) algorithm. Rejection sampling is used to ensure that all draws satisfy stability at each point in time.

4. $G(C_t \backslash \Phi_t, \tilde{h}_t, D_t, S)$. As in step 2 above, the VAR model in equation 5 can be written as $C_t \eta_t = D_t^{1/2} \bar{\eta}_t, \bar{\eta}_t \tilde{} N(0,1)$. Given a block diagonal S in transition equation 7, the Carter and Kohn (1994) algorithm can again be used to draw the the time-varying elements of C_t.

5. $G(D_t \backslash C_t, \Phi_t, \tilde{h}_t, q)$. Given C_t, Φ_t the orthogonal residuals of the VAR model in equation 5 can be calculated. A uni-variate stochastic volatility model applies to each residual with the transition equation . We draw each d_t using a multi-move step employing the recently developed

[1] As described in the appendix, the prior for is set in an identical fashion Q_Γ.

particle Gibbs algorithm with ancestoral sampling (see Andrieu *et al.* (2010), Lindsten *et al.* (2014)). This algorithm enables us to draw from the conditional posterior of the state-variable in a non-linear model without the need for linearisation (see Sangjoon Kim (1998)).

6. $G\left(\tilde{h}_t \backslash \Gamma_t, A_t, \Phi_t, Q_t\right)$. Equations 1 and 5 form a non-linear state-space system. As in step 5, we use a particle Gibbs step to sample from the conditional posterior of $\tilde{h}_t$.

7. The variances $Q_\Gamma, \breve{S}, Q_\Phi, S$ and q can be easily drawn from their respective conditional posterior distributions.

2.1.1 Estimation on artificial data

In order to test the algorithm, we conduct a small Monte-Carlo experiment. We generate artificial data from a bi-variate VAR with stochastic volatility featuring a change in the coefficients of the transition equation. In particular, the data generating process (DGP) is defined as

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix} = \begin{pmatrix} 0.7 & 0.1 \\ -0.1 & 0.7 \end{pmatrix} \begin{pmatrix} Z_{1t-1} \\ Z_{2t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0.1 & 1 \end{pmatrix} \begin{pmatrix} h_{1t} & 0 \\ 0 & h_{2t} \end{pmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$
$$\begin{pmatrix} \ln h_{1t} \\ \ln h_{2t} \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{pmatrix} \begin{pmatrix} \ln h_{1t-1} \\ \ln h_{2t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -0.1 & 1 \end{pmatrix} \begin{pmatrix} \exp(d_{1t}) & 0 \\ 0 & \exp(d_{2t}) \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}$$

where $\begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \tilde{} N(0, I_2)$ and $d_t = d_{t-1} + n_t, n_t \sim N(0, I_2 \times 0.001)$ with $d_t = \{d_{1t}, d_{2t}\}$. The coefficients $\begin{pmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{pmatrix}$ for the first half of the sample but change to $\begin{pmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{pmatrix} = \begin{pmatrix} 0.7 & -0.6 \\ 0.2 & 0.7 \end{pmatrix}$ for the second half implying a decline in the unconditional correlation between $\ln h_{1t}$ and $\ln h_{2t}$. We generate 600 observations, discarding the first 100 and using the remaining 500 for estimation. The model is estimated using 5000 iterations with a burn-in of 4000 and 20 particles employed in the particle Gibbs step. The experiment is repeated 200 times.

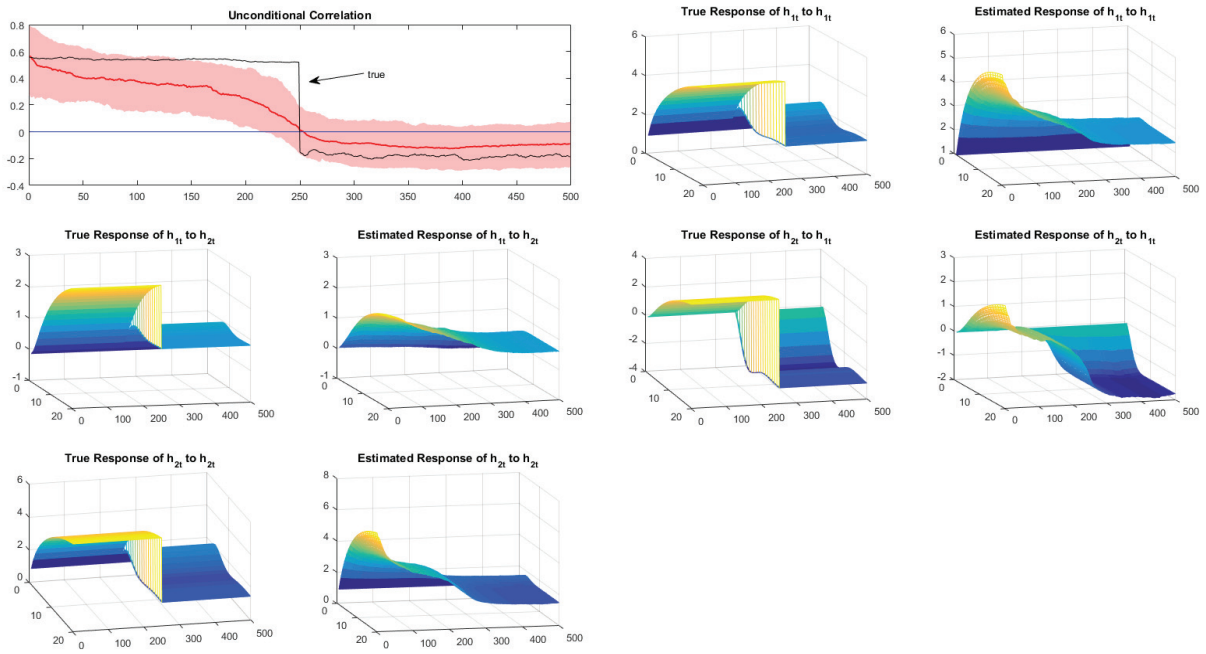

Figure 1: Estimation on artificial data. The figure reports true values and median estimates across the Monte-Carlo replications. The pink shaded area represents 1 standard deviation error band.

Figure 1 presents the results from this experiment. The top left panel shows the unconditional correlation between $\ln h_{1t}$ and $\ln h_{2t}$ calculated using an approximation to the unconditional covariance at each point in time (see equation 9). The black line which presents the correlation based on the DGP shows an abrupt shift at observation 250. The estimates from the proposed model display a smoother decline reflecting the law of motion embedded in the specification. However, the estimates provide a reasonable characterisation of the time-variation present in the correlation between $\ln h_{1t}$ and $\ln h_{2t}$. The remaining panels display the cumulated impulse responses derived using the transition equation for the stochastic volatilities. As expected, the median estimates display a more gradual shift than the DGP. However, as before the direction and the timing of the change provides a reasonable approximation of the underlying structural shift.

3 Empirical Analysis

The close relationship between the yield curve and the macroeconomy has been highlighted in several recent studies (see for e.g. Diebold *et al.* (2006)). Moreover, Mumtaz and Surico (2009) and Bianchi *et al.* (2009) show that this relationship is characterised by time-varying dynamics and heteroscedasticity. While the relationship between the level of government bond yields and macroeconomic variables has received substantial attention, the possibility of a link between the conditional variances associated with these variables has not been considered in the literature. From an economic perspective there are several reasons that suggest that such a link may exist and its strength may change over time. The volatility of the long-term interest rate is related closely to the term-premium (the wedge between the long-term interest rate and expectations about policy rates). The importance of the term-premium has been stressed by several studies (Wright (2011), Rudebusch and Swanson (2012), Swanson (2015)) and is considered as one of the main driver of the long-rate. From a policy perspective, the importance of the term premium has been enhanced further as 'unconventional' monetary policies are communicated in terms of long-rates (Bernanke (2013)). Thus an estimate of the time-varying correlation between long rate volatility and the volatility of macroeconomic variables provides valuable information on the possible impact of such policies. Similarly an estimate of the co-movement of long rate volatility across countries can enhance our understanding of the cross-country correlation of risk premia (see (Wright (2011) and Jotikasthira *et al.* (2015)) and the prevalance of risk sharing (see Colacito and Croce (2013), Kollmann (2015), Gourio *et al.* (2013) and Benigno *et al.* (2011)).

In order to explore this further, we estimate the TVP-VAR model proposed above using $Z_t = \{Y_t^{US}, R_t^{US}, Y_t^{UK}, R_t^{UK}\}$ where Y_t^i denotes real GDP growth in country i while R_t^i is the 10 year Government bond yield. This model, thus, not only allows us to investigate volatility comovement within the United States and the United Kingdom but also provides information on cross-country links in volatility as considered in Mumtaz and Theodoridis (2015).

The data on Z_t is quarterly and runs from 1955Q2 to 2015Q2. Note that we use the first ten years of data as a training sample and use the remaining years for estimation. US real GDP is obtained from the Federal Reserve Bank of St. Louis (FRED) database (Code: GDPC96). The US long term yield is downloaded from Global Financial data (Code: IGUSA10D). UK real GDP is obtained from the Office of National Statistics (Code: ABMI). The 10 year bond yield is obtained from the Bank of England's long-run database. The lag length is set to 4 in the benchmark model.

The model is estimated using 200,000 iterations of the MCMC algorithm described above. The first 100,000 iterations are discarded as burn-in and we retain every 20th draw from the remainder for inference. The appendix shows that the estimated inefficiency factors are fairly low providing some evidence for convergence.

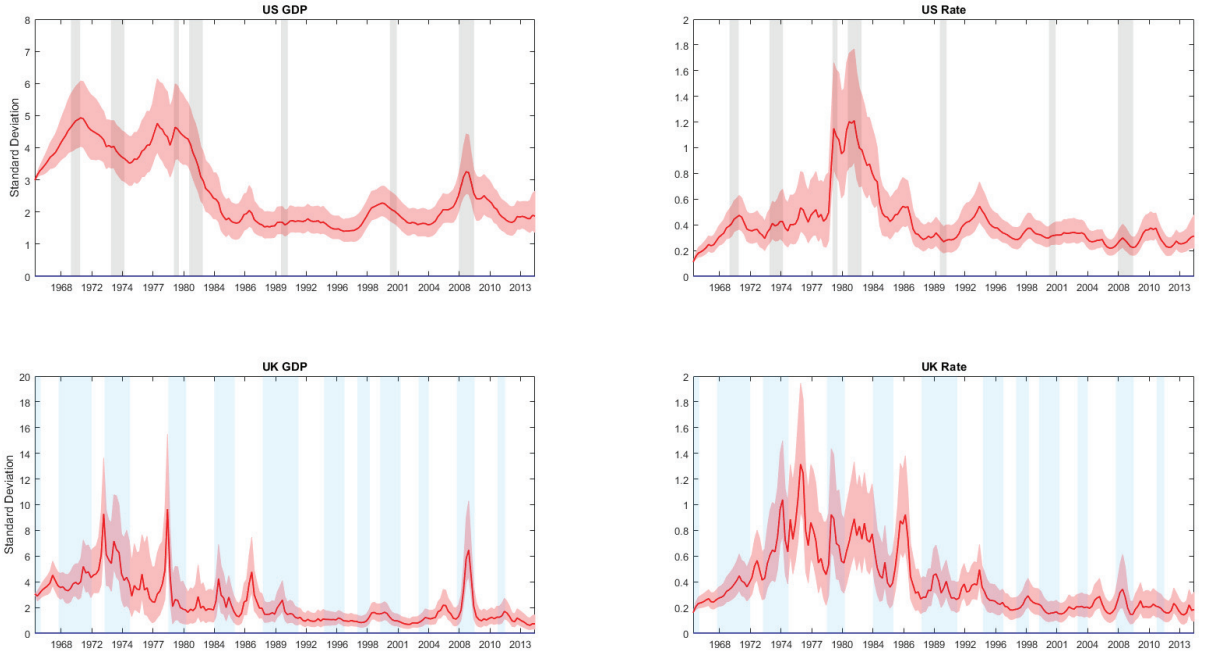

Figure 2: Stochastic volatility of the VAR shocks. The red sold lines are posterior medians while the red shaded area represents the 68% error bands. The grey and light blue vertical shaded areas represent US and UK recessions as reported by the NBER and OECD respectively.

3.1 Empirical results

We first compare the marginal likelihood (ML) of the proposed model to that of a standard TVP-SVOL which assumes an independent random walk specification as the transition equation for the stochastic volatilties. The ML is computed using the estimator proposed in Chib (1995) and described in Appendix B in context of the model proposed in this paper. The estimated log ML for the benchmark model is 9434.5. This is substantially larger than the estimate for the TVP-SVOL model: 4207.85. This provides strong evidence in favour the TVP-VAR that features time-varying dynamics in the transition equation for $\tilde{h}_t$.

Figure 2 presents the posterior estimates of the stochastic volatility, i.e. $\exp\left(\tilde{h}_t\right)^{1/2}$. Consider the results for the US presented in the top panel. The great moderation is clearly visible in the estimated volatility associated with the GDP shock – the shock standard deviation declined dramatically during the early 1980s and remained relatively low until the 'great recession' in 2008. The shock volatility of the US long-term rate was at its peak during the late 1970s and the early 1980s, with this episode coinciding with the Volcker experiment of targeting non-borrowed reserves. This volatility declined subsequently to levels seen in the earlier part of the sample. The bottom left panel of the figure shows that the shock variance associated with UK GDP growth remained high until the early 1990s. The post-1992 inflation-targeting period was characterised by stability until 2008, when this shock volatility spiralled to the levels seen in the 1970s and the 1980s. The evolution of the standard deviation of the UK long-rate shock displays a similar pattern–the volatility is high pre-1992 followed by a stable period until the recent financial crisis. In general, the temporal evolution of the stochastic volatilities is estimated to be similar to that reported in previous studies employing TVP-SVOL models (see Cogley and Sargent (2005) and Benati (2008)). In contrast to these previous studies, however, we are able to explore if there was a systematic co-movement between the elements of $\tilde{h}_t$. We turn to this analysis next.

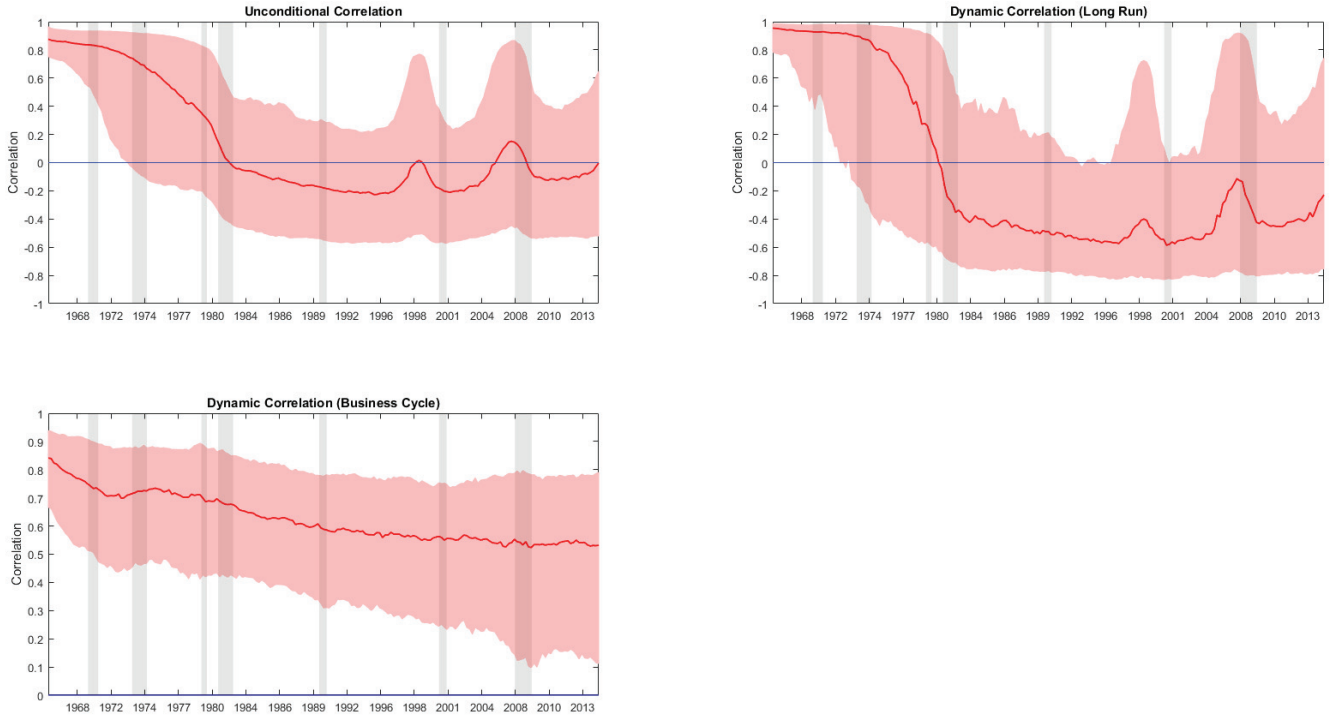

Figure 3: Correlation between the conditional volatility of US GDP growth and the volatility of the US long-term interest rate. The red sold lines are posterior medians while the red shaded area represents the 68% error bands. The grey vertical shaded areas represent US recessions as reported by the NBER.

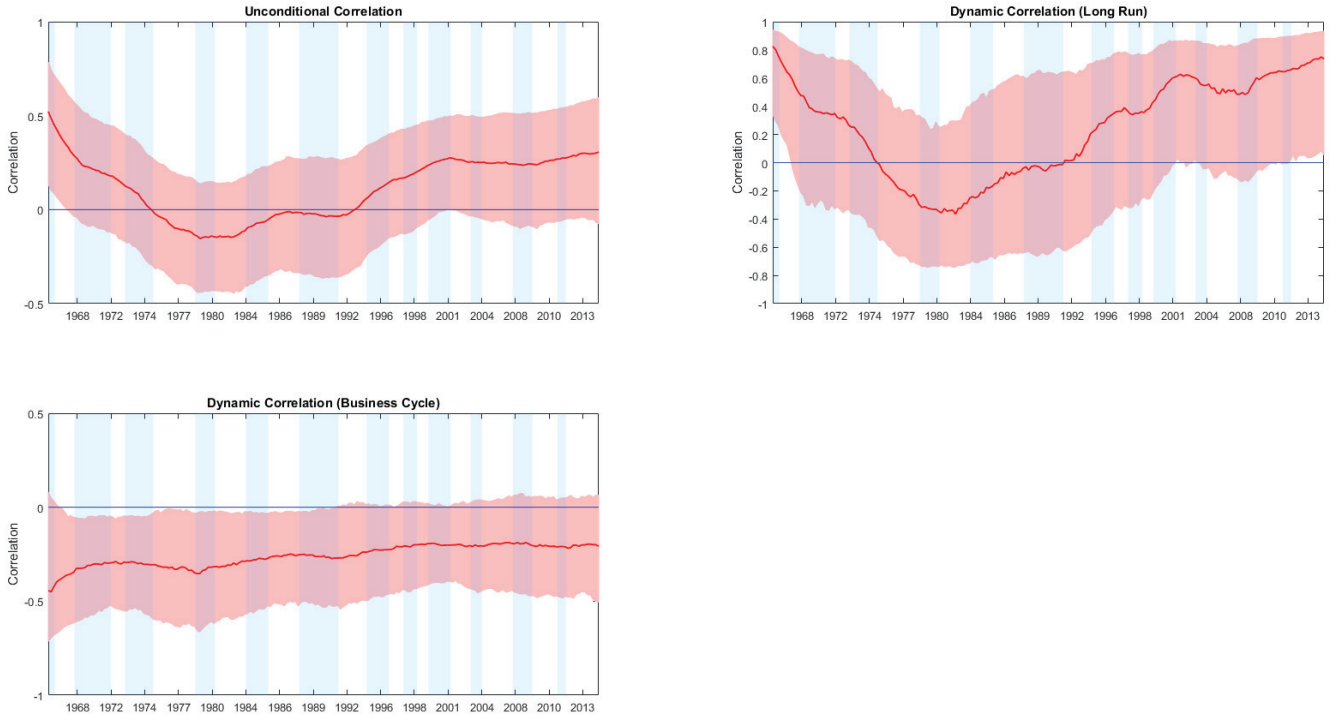

Figure 4: Correlation between the conditional volatility of UK GDP growth and the volatility of the UK long-term interest rate. The red sold lines are posterior medians while the red shaded area represents the 68% error bands. The blue vertical shaded areas represent UK recessions as reported by the OECD.

3.1.1 Dynamic Correlations

Note that an approximation to the unconditional variance-covariance of $\tilde{h}_t$ can be calculated at each point in time using the parameters of equation 5:

$$(I - \theta_t \otimes \theta_t)^{-1} vec\,(Q_t) \tag{9}$$

The time-varying correlation matrix can easily be computed as a by-product. Similarly, the spectral density matrix of $\tilde{h}_t$ can be calculated at each point in time as:

$$\hat{f}_t(\omega) = (I - \theta_t e^{-i\omega})^{-1} \frac{Q_t}{2\pi} \left[(I - \theta_t e^{-i\omega})^{-1}\right]', \tag{10}$$

where ω denotes the frequency. The off-diagonal elements of the spectral density matrix summarises the relationship between $\tilde{h}_t$ at different frequencies. We focus on a particular measure of association called dynamic correlations proposed in Christophe Croux (2001). This measure is defined as:

$$\frac{\hat{c}_{ij}\,(\omega)}{\sqrt{\hat{f}_t^{ii}(\omega)\hat{f}_t^{jj}(\omega)}}, \tag{11}$$

where $\hat{c}_{ij}\,(\omega)$ denotes the cospectrum between the i_{th} and j_{th} volatility at frequency ω. The dynamic correlation lies between -1 and 1. It equals one if $\tilde{h}_{it}$ and $\tilde{h}_{jt}$ are exactly synchronised at a given frequency.

Figure 3 shows the estimated correlation between the stochastic volatilities associated with the US GDP shock and the US long-rate shock. The first panel shows the unconditional correlation. The remaining panels show the dynamic correlation at the long-run and the business cycle frequencies. The long run frequency corresponds to cycles of 100 years, while the business cycle frequency is associated with cycles of 3 years. The figure shows that the correlation between these conditional volatilties changed dramatically over the sample period. In the pre-1980 period, this correlation was close to 1. However the onset of the great moderation coincided with a sharp decline in this volatility co-movement. After the mid-1980s the median correlation remains largely below zero, with the period leading up to great recession providing one exception. The estimated dynamic correlations suggest that this decline in co-movement was concentrated at the long-run frequencies. While, the business cycle association between these conditional volatilities did decline over time, the magnitude of the decline was much less dramatic. The evolution of this correlation followed a similar pattern during the 1970s in the UK (figure 4). That is, the volatilities of the output growth and the long rate shock were characterised by high co-movement which then declined as the decade progressed. In contrast to the US, however, the estimated median correlation for the UK displays an increase after 1992. The top right panel of the figure suggests that this increase largely occurs at the long run frequency with the dynamic correlation reaching its pre-1980 level in the period after the Bank of England was granted operational independence.

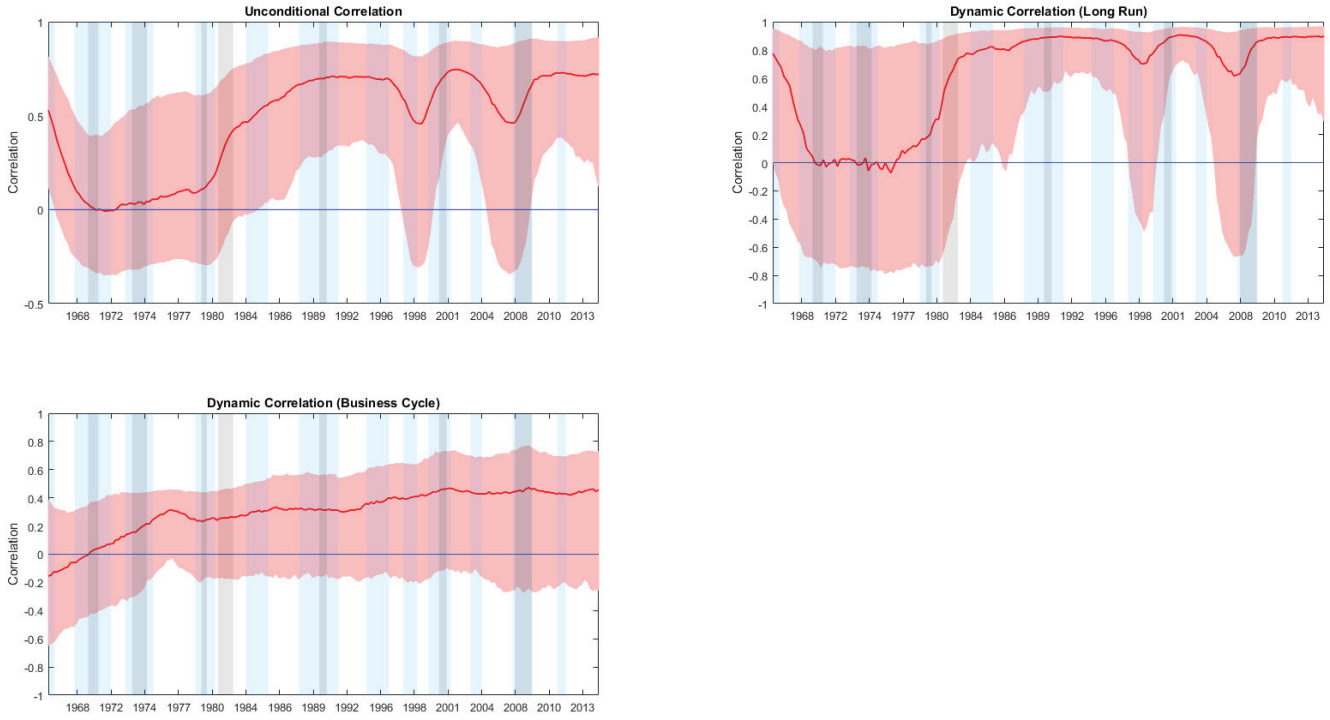

Figure 5: Correlation between the conditional volatility of US GDP growth and the volatility of the UK GDP Growth. The red sold lines are posterior medians while the red shaded area represents the 68% error bands. The grey and light blue vertical shaded areas represent US and UK recessions as reported by the NBER and OECD respectively.

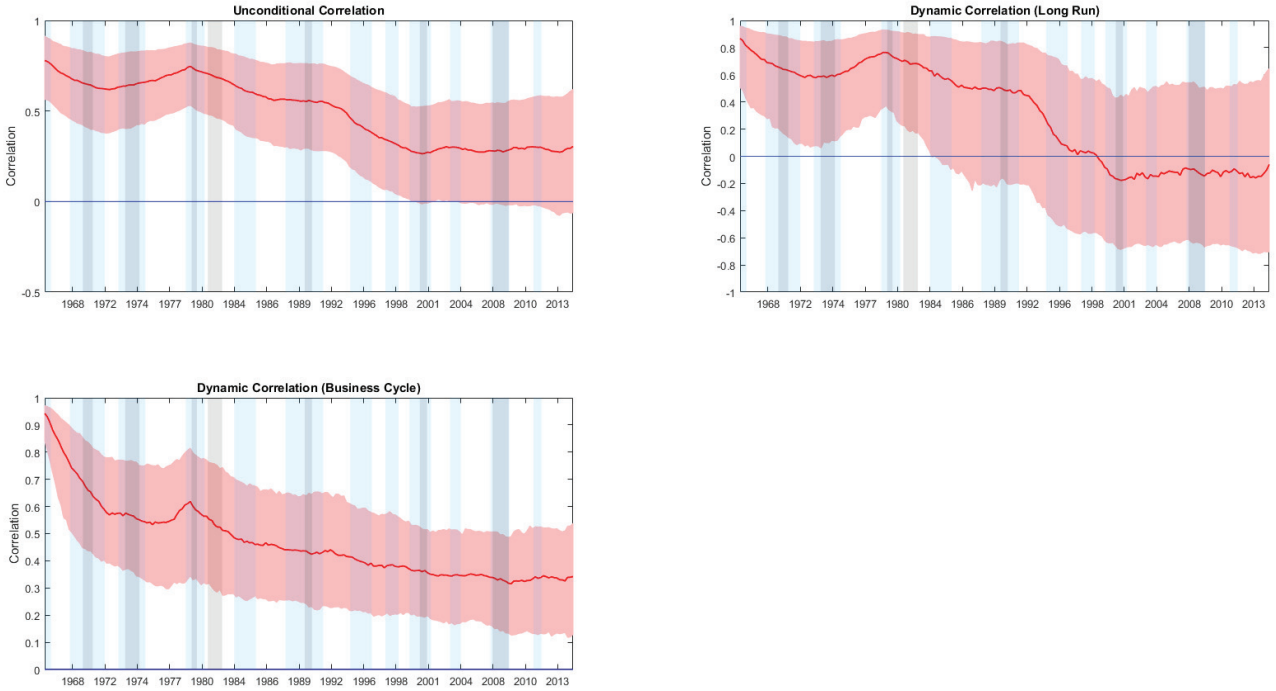

Figure 6: Correlation between the conditional volatility of US long-term rate and the volatility of the UK long-term rate. The red sold lines are posterior medians while the red shaded area represents the 68% error bands. The grey and light blue vertical shaded areas represent US and UK recessions as reported by the NBER and OECD respectively.

Figures 5 and 6 consider the correlation between the conditional volatilities across countries. Figure 5 shows that the correlation between GDP growth volatilties increased dramatically after the mid-1980s with long run correlation close to 1 over the post-1985 period. This supports the results in Mumtaz and Theodoridis (2015) who also show an increasing co-movement in the volatility of output growth across countries. In contrast, figure 6 shows that the conditional variance of the long-term rate has become less correlated across these two countries. After the late 1990s, the unconditional correlation is about half of the estimate of the early 1980s and the long-run dynamic correlation is estimated to be negative.

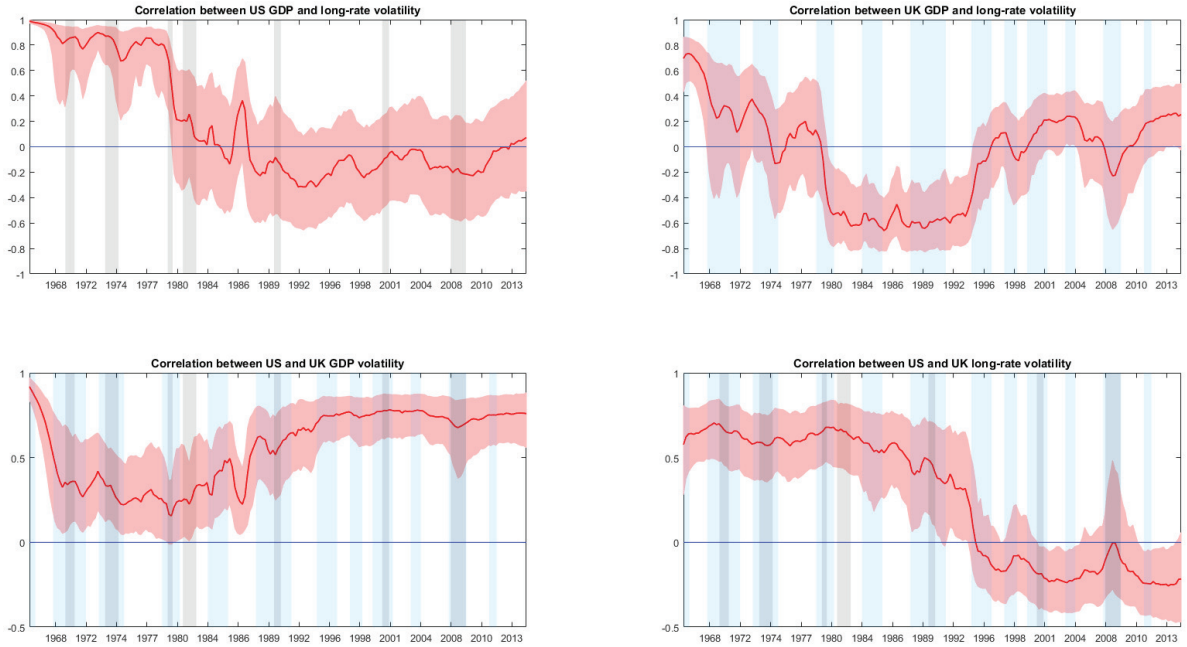

Figure 7: Unconditional correlations between volatilities from a model using alternative priors for Q_Φ. The grey and light blue vertical shaded areas represent US and UK recessions as reported by the NBER and OECD respectively.

To check the robustness we explore if the choice of prior for Q_Φ affects the degree of time-variation estimated in the correlations. In particular, we re-estimate the benchmark model setting the scaling factor κ equal to a reduced value of $1e-04$. The degrees of freedom remain as in the benchmark model. Thus we incorporate a prior belief favouring a smaller degree of time-variation. The estimated unconditional correlations from this alternative model are presented in figure 7. It is immediately clear that these are very similar to the benchmark case suggesting therefore that the key results are not driven by the choice of the prior for Q_Φ.

In summary, this analysis suggests three main conclusions. First, the correlation between the conditional variance of US output growth and the conditional variance of the long-term rate was high during the 1970s and the early 1980s but declined substantially thereafter. A similar decline is evident for the UK, but the correlation appears to have increased again over the inflation targeting period. Second, the correlation between the conditional variance of US and UK GDP shows a dramatic increase over the great moderation period. Finally, the co-movement between the volatility of the long-rate in these two countries was strong until the early 1990s–over the inflation targeting period in the UK, however, the null hypothesis of zero correlation between these variances cannot be rejected.

3.1.2 Discussion

The discussion in this section sketches out some of the structural changes and transmission channels that they could be consistent with the estimates presented above.[2] The decline in the correlation between GDP growth and long-term rates volatility seems to be consistent with the more weight placed by monetary authorities on the stabilisation of inflation. As authorities target inflation more aggressively, inflation premia (perhaps one of the most important component of the term-premium) decreases as market participants understand that the likelihood that long-run inflation expectation will be away from their target is small. In this case they do not demand a premium on investing on long-term debt as it is unlikely that high inflation is going to erode their investment cash flows. Assuming that the volatility of GDP growth has a real and nominal component, the decrease in the term-premium could be consistent with a fall in nominal volatility in the economy and lead to a lower correlation between the variance of GDP and the long rate.[3]

The increase in the correlation between US and UK GDP growth volatility could be consistent with more global markets and the reduced ability of domestic agents to share risks with the rest of the world (Colacito and Croce (2013) and Campbell (2005)). For instance, as the degree of home bias decreases, the contribution of imports (net trade) to GDP rises. In this situation domestic households and firms are unable to diversify risks that originate abroad making it more likely that domestic volatility is going to co-move with the foreign one.

In order to understand the declining correlation between US and UK long-term interest rate volatility, it is important to take the exchange rate regime into account as well. It is perhaps not a coincidence that the fall in the correlation estimated by the model takes place around 1992 when the sterling was withdrawn from the European Exchange Rate Mechanism ('Black Wednesday'), which marks the beginning of the UK floating exchange regime. As explained in Obstfeld and Taylor (2003), monetary authorities cannot achieve a stable exchange rate, perfect capital mobility and pursue an independent monetary policy simultaneously. Thus, prior to 1992, given a managed exchange rate it is likely that long run domestic inflation expectations closely followed those prevalent in major foreign economies. In such an environment it is not difficult to understand that if

[2] The discussion in this section is based on simulations from the model developed in Chin *et al.* (2015).

[3] The rebound in this correlation in the UK would suggest that real volatility also declined significantly after the introduction of inflation targeting.

foreign investors demand a compensation (inflation premium) to invest in foreign long-term debt due to the volatility of the inflation, then it is very likely that domestic investors will request a similar compensation to invest in domestic long-run debt as they are equally exposed to inflation risk. After the ERM crisis and the introduction of inflation targeting, long run inflation expectations returned to target in the UK with domestic market participants willing to undertake long term investments without requiring an additional inflation premium. Such a change would be consistent with a fall in the correlation of long-rate volatility across the two countries.

4 Conclusion

In this paper we extend the TVP-SVOL model to allow for a time-varying dynamic relationship among the stochastic volatilties. This extended model allows researchers to investigate changes in the relationship between conditional variances. Such changes are consistent with structural shifts seen recently in the industrialised countries. Using this model, we show that the co-movement between the conditional variance of GDP growth and the long-term rate has become weaker over time in the US. Similarly, the correlation between the variance of the long-term interest rate across the US and the UK declined over the great moderation period. In contrast, the volatility of US and UK GDP appears to have become increasingly correlated. We argue that these estimates are consistent with (a) a decline in the term-premium possibly resulting from policy-makers focus on inflation control and (b) a greater degree of globalisation.

In future work, it would be useful to explore whether a DSGE model featuring time-varying volatility can be used to replicate some of these second moment co-movements. It may also be interesting to investigate if the relationship between the conditional variance of real variables and financial series such as stock returns and credit spreads has remained stable over time. This would be provide additional insights on the impact of the great moderation and the consequences of the great recession.

References

Andrieu, Christophe, Arnaud Doucet and Roman Holenstein, 2010, Particle Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society Series B* **72**(3), 269–342.

Arulampalam, M. Sanjeev, Simon Maskell and Neil Gordon, 2002, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Transactions on Signal Processing* **50**, 174–188.

Benati, Luca, 2008, The Great Moderation in the United Kingdom, *Journal of Money, Credit and Banking* **40**(1), 121–47.

Benati, Luca and Haroon Mumtaz, 2007, U.S. evolving macroeconomic dynamics - a structural investigation, *Working Paper Series 746*, European Central Bank.

Benigno, Gianluca, Pierpaolo Benigno and Salvatore Nisticï£¡, 2011, Risk, Monetary Policy and the Exchange Rate, *in* Daron Acemoglu and Michael Woodford (editors), *NBER Macroeconomics Annual 2011, Volume 26*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 247–309.

Berger, Tino, Sibylle Herz and Bernd Kempa, 2014, Global Macroeconomic Uncertainty, *mimeo*, University of Bern.

Bernanke, Ben S., 2013, Long-Term Interest Rates, Speech at the Annual Monetary/Macroeconomics Conference: The Past and Future of Monetary Policy.

Bianchi, Francesco, Haroon Mumtaz and Paolo Surico, 2009, The great moderation of the term structure of UK interest rates, *Journal of Monetary Economics* **56**(6), 856 – 871.

Campbell, Sean D., 2005, Stock market volatility and the Great Moderation, *Technical report*.

Canova, Fabio and Fernando J. Perez Forero, 2015, Estimating overidentified, nonrecursive, time-varying coefficients structural vector autoregressions, *Quantitative Economics* **6**(2), 359–384.

Carter, C and P Kohn, 1994, On Gibbs sampling for state space models, *Biometrika* **81**, 541–53.

Chib, Siddhartha, 1995, Marginal Likelihood from the Gibbs Output, *Journal of the American Statistical Association* **90**(432), 1313–1321.

Chin, Michael, Thomai Filippeli and Konstantinos Theodoridis, 2015, Cross-country co-movement in long-term interest rates: a DSGE approach, *Bank of England working papers 530*, Bank of England.

Christophe Croux, Mario Forni, Lucrezia Reichlin, 2001, A Measure of Comovement for Economic Variables: Theory and Empirics, *The Review of Economics and Statistics* **83**(2), 232–241.

Cogley, T. and T. J. Sargent, 2005, Drifts and Volatilities: monetary policies and outcomes in the Post WWII U.S., *Review of Economic Dynamics* **8**, 262–302.

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

Colacito, Riccardo and Mariano M. Croce, 2013, International Asset Pricing with Recursive Preferences, *Journal of Finance* **68**(6), 2651–2686.

Diebold, Francis X., Glenn D. Rudebusch and S. BoragËĞan Aruoba, 2006, The macroeconomy and the yield curve: a dynamic latent factor approach, *Journal of Econometrics* **131**(1âĂŞ2), 309 – 338.

Fuentes-Albero, Cristina, 2012, Financial Frictions, Financial Shocks, and Aggregate Volatility, *Dynare Working Papers 18*, CEPREMAP.

Gourio, Francois, Michael Siemer and Adrien Verdelhan, 2013, International risk cycles, *Journal of International Economics* **89**(2), 471 – 484.

Jotikasthira, Chotibhak, Anh Le and Christian Lundblad, 2015, Why do term structures in different currencies co-move?, *Journal of Financial Economics* **115**(1), 58 – 83.

Kim, Chang-Jin and Charles R. Nelson, 1999, Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle, *The Review of Economics and Statistics* **81**(4), 608–16.

Kollmann, Robert, 2015, Exchange Rates Dynamics with Long-Run Risk and Recursive Preferences, *Open Economies Review* **26**(2), 175–196.

Lindsten, Fredrik, Michael I. Jordan and Thomas B. Schön, 2014, Particle Gibbs with Ancestor Sampling, *Journal of Machine Learning Research* **15**, 2145–2184.

Lopes, Hedibert Freitas and Carlos Marinho Carvalho, 2007, Factor stochastic volatility with time varying loadings and Markov switching regimes, *Journal of Statistical Planning and Inference* **137**(10), 3082 – 3091. Special Issue: Bayesian Inference for Stochastic Processes.

Mike K. P. So, K. Lam, W. K. Li, 1998, A Stochastic Volatility Model with Markov Switching, *Journal of Business Economic Statistics* **16**(2), 244–253.

Mumtaz, Haroon and Konstantinos Theodoridis, 2015, Common and Country Specific Economic Uncertainty, *Working Papers 752*, Queen Mary University of London, School of Economics and Finance.

Mumtaz, Haroon and Laura Sunder-Plassmann, 2013, Time-Varying Dynamics Of The Real Exchange Rate: An Empirical Analysis, *Journal of Applied Econometrics* **28**(3), 498–525.

Mumtaz, Haroon and Paolo Surico, 2009, Time-varying yield curve dynamics and monetary policy, *Journal of Applied Econometrics* **24**(6), 895–913.

Nonejad, Nima, 2015, REPLICATING THE RESULTS IN ŚA NEW MODEL OF TREND INFLATIONŠ USING PARTICLE MARKOV CHAIN MONTE CARLO, *Journal of Applied Econometrics* pp. n/a–n/a.

Obstfeld, Maurice and Alan M. Taylor, 2003, Globalization and Capital Markets, *Globalization in Historical Perspective*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 121–188.

Prieto, Esteban, Sandra Eickmeier and Massimiliano Marcellino, 2016, Time Variation in Macro-Financial Linkages, *Journal of Applied Econometrics* pp. n/a–n/a.

Primiceri, G, 2005, Time varying structural vector autoregressions and monetary policy, *The Review of Economic Studies* **72**(3), 821–52.

Rudebusch, Glenn D. and Eric T. Swanson, 2012, The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks, *American Economic Journal: Macroeconomics* **4**(1), 105–43.

Sangjoon Kim, Neil Shephard, Siddhartha Chib, 1998, Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models, *The Review of Economic Studies* **65**(3), 361–393.

Swanson, Eric, 2015, A Macroeconomic Model of Equities and Real, Nominal, and Defaultable Debt, *Mimeo*, University of California Irvine.

Wright, Jonathan H., 2011, Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset, *American Economic Review* **101**(4), 1514–34.

A Appendix A: Model Estimation

Consider the VAR model:

$$
\begin{aligned}
Z_t &= c_t + \sum_{j=1}^{P} \beta_t Z_{t-j} + \Omega_t^{1/2} e_t, e_t \sim N(0,1) && (12)\\
\Gamma_t &= vec\left([\beta_t, c_t]\right) && (13)\\
\Gamma_t &= \Gamma_{t-1} + \tau_t, \tau_t \tilde{} N\left(0, Q_\Gamma\right) && (14)\\
\Omega_t &= A_t^{-1} H_t A_t^{-1'} && (15)\\
a_t &= a_{t-1} + \bar{s}_t, VAR(\bar{s}_t) = \breve{S} && (16)\\
\tilde{h}_t &= \alpha_t + \theta_t \tilde{h}_{t-1} + \eta_t, \eta_t \sim N(0, Q_t), E\left(e_t, \eta_t\right) = 0 && (17)\\
\Phi_t &= vec([\theta_t, \alpha_t]) && (18)\\
\Phi_t &= \Phi_{t-1} + u_t, u_t \tilde{} N(0, Q_\Phi) && (19)\\
Q_t &= C_t^{-1} D_t C_t^{-1'} && (20)\\
c_t &= c_{t-1} + s_t, VAR(s_t) = S && (21)\\
d_t &= d_{t-1} + n_t, n_t \sim N(0, q) && (22)
\end{aligned}
$$

where $\tilde{h}_t = [h_{1t}, h_{2t}, ..h_{N,t}]$, $H_t = diag\left(\exp\left(\tilde{h}_t\right)\right)$. Similarly $d_t = [d_{1t}, d_{2t}, ..d_{N,t}]$, $D_t = diag\left(\exp\left(d_t\right)\right)$.

A.1 Prior distributions and starting values

A.1.1 VAR Coefficients

The initial conditions for the VAR coefficients Γ_0 in equation 12 are obtained via an OLS estimate of a fixed coefficient VAR using the first $T_0 = 40$ observations of the sample period. Let Γ^{ols} and $\hat{V}^{ols}$ denote the OLS estimate of the VAR coefficients and the covariance matrix estimated on the pre-sample data. The prior for $\Gamma_0 \tilde{} N(\Gamma^{ols}, var(\Gamma^{ols}))$. The prior on Q_Γ is assumed to be inverse Wishart $IW\left(\bar{Q}_\Gamma, TT_0\right)$ where $\bar{Q}_\Gamma$ is assumed to be $TT_0 \times var(\Gamma^{ols}) \times k$ and T_0 is the length of the sample used to for calibration and TT_0 equals the rows of Q_Γ plus 10. Following Cogley and Sargent (2005), the scaling factor k is set to 3.5×10^{-4} in the benchmark case.

A.1.2 Elements of A

The prior for the off-diagonal elements A_t is $A_0 \sim N\left(a^{ols}, V\left(a^{ols}\right)\right)$ where a^{ols} are the off-diagonal elements of $\hat{V}^{ols}$, with each row scaled by the corresponding element on the diagonal. $V\left(a^{ols}\right)$ is assumed to be diagonal with the elements set equal to 10 times the absolute value of the corresponding element of a^{ols}. The prior distribution for the blocks of $\breve{S}$ is inverse Wishart: $\breve{S}_{i,0} \sim IW(\breve{S}_i, K_i)$ where $i = 1..N-1$ indexes the blocks of $\breve{S}$. $\breve{S}_i$ is calibrated using a^{ols}. Specifically, $\breve{S}_i$ is a diagonal matrix with the diagonal elements given by 10^{-3}. A similar prior specification is used in previous studies such as Benati and Mumtaz (2007).

A.1.3 Elements of H_t

Following Cogley and Sargent (2005) we use the training sample to set the prior for the elements of the transition equation of the model. Let $\hat{V}^{ols}$ denote the OLS estimate of the VAR covariance matrix estimated on the pre-sample data of $T_0 = 40$ observations. The prior for $\tilde{h}_t$ at $t = 0$ is defined as $\ln h_0 \sim N(\ln \mu_0, I_4)$ where μ_0 are the diagonal elements of the Cholesky decomposition of $\hat{V}^{ols}$. To obtain starting values for h_t we estimate a VAR model that allows for a univariate stochastic volatility process for each orthogonalised VAR shock. Denote this starting value by $\tilde{h}_t^0$.

A.1.4 Coefficients of the transition equation

The initial conditions for the VAR coefficients Φ_0 (eq 17) are obtained via an OLS estimate of a fixed coefficient VAR using the first $T_0 = 40$ observations of the sample period. The VAR is estimated using the starting value $\ln \tilde{h}_t^0$. Let Φ^{ols} and $\hat{v}^{ols}$ denote the OLS estimate of the VAR coefficients and the covariance matrix estimated on the pre-sample data described above. The prior for $\Phi_0 \tilde{} N(\Phi^{ols}, var(\Phi^{ols}))$. The prior on Q_Φ is assumed to be inverse Wishart $Q_{\Phi,0} \sim IW\left(\bar{Q}_{\Phi,0}, TT_0\right)$ where $\bar{Q}_{\Phi,0}$ is assumed to be $TT_0 \times var(\Phi^{ols}) \times k$ and T_0 is the length of the sample used to for calibration and TT_0 equals the rows of Q_Φ plus 10. Following Cogley and Sargent (2005), the scaling factor k is set to 3.5×10^{-4} for the benchmark case.

A.1.5 Elements of C_t

The prior for the off-diagonal elements C_t is $C_0 \sim N\left(c^{ols}, V\left(c^{ols}\right)\right)$ where c^{ols} are the off-diagonal elements of $\hat{v}^{ols}$, with each row scaled by the corresponding element on the diagonal. $V\left(c^{ols}\right)$ is assumed to be diagonal with the elements set equal to 10 times the absolute value of the corresponding element of c^{ols}. The prior distribution for the blocks of S is inverse Wishart: $S_{i,0} \sim IW(\bar{S}_i, K_i)$ where $i = 1..N-1$ indexes the blocks of S. $\bar{S}_i$ is calibrated using c^{ols}. Specifically, $\bar{S}_i$ is a diagonal matrix with diagonal elements given by 10^{-3}. This prior specification is used in previous studies such as Benati and Mumtaz (2007).

A.1.6 Elements of D_t

Following Cogley and Sargent (2005) we use the training sample described above to set the prior for the elements of the transition equation of the model. The prior for d_t at $t = 0$ is defined as $\ln d_0 \sim N(\ln \mu_{d0}, I_4)$ where μ_{d0} are the diagonal elements of the Cholesky decomposition of the error covariance matrix obtained by estimating a fixed coefficient VAR using $\tilde{h}_t^0$ as the endogenous variables. The prior for the elements of q is inverse Gamma: $p(q) \tilde{} IG(q_0, vq_0)$. The degrees of freedom are set to 5. The scale parameters q_0 are calibrated by estimating a VAR on the starting values $\tilde{h}_t^0$ that features univariate stochastic volatility models for the orthogonalised residuals.

A.2 Simulating the posterior distributions

The MCMC algorithm samples from the following conditional posterior distributions.

A.2.1 VAR coefficients $H\left(\Gamma_t | Q_\Gamma, A_t, H_t, \breve{S}, \Phi_t, Q_\Phi, C_t, D_t, S, q\right)$

Given a draw for the stochastic volatility $\tilde{h}_t$ and the covariance matrix Q_Γ the VAR model represents a linear Gaussian state space mode. The Carter and Kohn (1994) algorithm can be applied to draw from the conditional posterior of Γ_t. The distribution of the time-varying coefficients conditional on all other parameters Ξ is linear and Gaussian: $\Gamma_t \backslash \tilde{h}_t, \Xi \sim N\left(\Gamma_{T\backslash T}, P_{T\backslash T}\right)$ and $\Gamma_t \backslash \Gamma_{t+1}, \tilde{h}_t, \Xi \sim N\left(\Gamma_{t\backslash t+1,\Gamma_{t+1}}, P_{t\backslash t+1,\Gamma_{t+1}}\right)$ where $t = T-1, ..1$. As shown by Carter and Kohn (1994) the simulation proceeds as follows. First we use the Kalman filter to draw $\Gamma_{T\backslash T}$ and $P_{T\backslash T}$ and then proceed backwards in time using $\Gamma_{t|t+1,\Gamma_{t+1}} = \Gamma_{t|t} + P_{t|t} P_{t+1|t}^{-1}\left(\Gamma_{t+1} - \Gamma_{t|t}\right)$ and $P_{t|t+1,\Gamma_{t+1}} = P_{t|t} - P_{t|t} P_{t+1|t}^{-1} P_{t|t}$.

A.2.2 Covariance Matrix Q_Γ $H\left(Q_\Gamma|\Gamma_t, A_t, H_t, \breve{S}, \Phi_t, Q_\Phi, C_t, D_t, S, q\right)$

Given a draw of Γ_t, the covariance matrix Q_Γ can be drawn from the inverse Wishart density with scale matrix $\tau_t'\tau_t + \bar{Q}_{\Phi,0}$ and degrees of freedom $T + TT_0$

A.2.3 Element of A_t $H\left(A_t|Q_\Gamma, \Gamma_t, H_t, \breve{S}, \Phi_t, Q_\Phi, C_t, D_t, S, q\right)$

Given a draw for Γ_t and H_t the VAR model can be written as $A_t'\left(\tilde{Z}_t\right) = \tilde{e}_t$ where $\tilde{Z}_t = Z_t - c_t - \sum_{j=1}^{P}\beta_t Z_{t-j}$. This is a system of linear equations with time-varying coefficients and a known form of heteroscedasticity. The jth equation of this system is given as $\tilde{Z}_{jt} = a_{jt}\tilde{Z}_{-jt} + \tilde{e}_{jt}$ where the subscript j denotes the jth column of $\tilde{Z}_t$ while $-j$ denotes columns 1 to $j-1$. Note that the variance of $\tilde{e}_{jt}$ are time-varying and given by $exp(h_{jt})$. The time-varying coefficient follows the process $a_{jt} = a_{jt-1} + \bar{s}_{jt}$ with the shocks to the jth equation $\bar{s}_{jt}$ uncorrelated with those from other equations. In other words the covariance matrix is assumed to be block diagonal as in Primiceri (2005). With this assumption in place, the Carter and Kohn (1994) algorithm can be applied to draw the time varying coefficients for each equation of this system seperately.

A.2.4 Covariance Matrix $\breve{S}$ $H\left(\breve{S}|Q_\Gamma, \Gamma_t, A_t, H_t, \Phi_t, Q_\Phi, C_t, D_t, S, q\right)$

Given a draw of A_t the conditional posterior for the blocks of $\breve{S}$ is inverse Wishart and these covariance matrices can be drawn easily. The scale matrix is $\bar{s}_{jt}'\bar{s}_{jt} + \breve{S}_{j,0}$ and the degrees of freedom are the sample size plus the prior degrees of freedom.

A.2.5 Coefficients of the transition equation Φ_t $H\left(\Phi_t|\breve{S}, Q_\Gamma, \Gamma_t, A_t, H_t, Q_\Phi, C_t, D_t, S, q\right)$

Given a draw for $\tilde{h}_t$ and Q_t, the transition equation is a VAR with time-varying coefficients. As this is a linear and Gaussian state space model, the Carter and Kohn (1994) algorithm can be applied to draw from the conditional posterior of Φ_t. The distribution of the time-varying loadings conditional on all other parameters is linear and Gaussian: $\Phi_t\backslash\tilde{h}_t, \Xi \sim N\left(\Phi_{T\backslash T}, \bar{P}_{T\backslash T}\right)$ and $\Phi_t\backslash\Phi_{t+1}, \tilde{h}_t, \Xi \sim N\left(\Phi_{t\backslash t+1,\Phi_{t+1}}, \bar{P}_{t\backslash t+1,\Phi_{t+1}}\right)$ where $t = T-1, ..1$, Ξ denotes a vector that holds all the other parameters and states of the transition equation. As shown by Carter and Kohn (1994) the simulation proceeds as follows. First we use the Kalman filter to draw $\Phi_{T\backslash T}$ and $\bar{P}_{T\backslash T}$ and then proceed backwards in time using $\Phi_{t|t+1,\Phi_{t+1}} = \Phi_{t|t} + \bar{P}_{t|t}\bar{P}_{t+1|t}^{-1}\left(\Phi_{t+1} - \Phi_{t|t}\right)$ and $\bar{P}_{t|t+1,\Phi_{t+1}} = \bar{P}_{t|t} - \bar{P}_{t|t}\bar{P}_{t+1|t}^{-1}\bar{P}_{t|t}$.

A.2.6 Covariance Matrix Q_Φ $H\left(Q_\Phi|\Gamma_t, A_t, H_t, \breve{S}, \Phi_t, Q_\Gamma, C_t, D_t, S, q\right)$

Given a draw of Φ_t, the covariance matrix Q_Φ can be drawn from the inverse Wishart density with scale matrix $u_t'u_t + \bar{Q}_{\Phi,0}$ and degrees of freedom $T + TT_0$

A.2.7 Elements of C_t $H\left(C_t|Q_\Gamma, \Gamma_t, H_t, \breve{S}, \Phi_t, Q_\Phi, A_t, D_t, S, q\right)$

Given a draw for the coefficients of the transition equation and stochastic volatility d_t the transition equation can be written as $C_t\left(\eta_t\right) = \bar{e}_t$ where η_t denotes the residuals. This is a system of linear equations with time-varying coefficients and a known form of heteroscedasticity. The jth equation of this system is given as $\eta_{jt} = c_{jt}\eta_{-jt} + \bar{e}_{jt}$ where the subscript j denotes the jth column of v while

$-j$ denotes columns 1 to $j-1$. Note that the variance of $\bar{e}_{jt}$ is time-varying and given by $\exp(d_{jt})$. The time-varying coefficient follows the process $c_{jt} = c_{jt-1} + s_{jt}$ with the shocks to the jth equation s_{jt} uncorrelated with those from other equations. In other words the covariance matrix is assumed to be block diagonal as in Primiceri (2005). With this assumption in place, the Carter and Kohn (1994) algorithm can be applied to draw the time varying coefficients for each equation of this system seperately.

A.3 Covariance Matrix $S\ H\left(S|Q_\Gamma, \Gamma_t, A_t, H_t, \Phi_t, Q_\Phi, C_t, D_t, \breve{S}, q\right)$

Given a draw of C_t the conditional posterior for the blocks of S is inverse Wishart and these covariance matrices can be drawn easily. The scale matrix is $s'_{jt}s_{jt} + S_{j,0}$ and the degrees of freedom are the sample size plus the prior degrees of freedom.

A.3.1 Elements of $D_t\ H\left(D_t|C_t, Q_\Gamma, \Gamma_t, H_t, \breve{S}, \Phi_t, Q_\Phi, A_t, S, q\right)$

Given a draw for the coefficients of the transition equation and C_t the orthogonalised resduals $C_t(\eta_t) = \bar{e}_t$ can be calculated. The jth residual follows a univariate stochastic volatility model

$$\begin{aligned} \bar{e}_{jt} &= \sqrt{\exp(d_{jt})}\bar{e}^*_{jt}, \bar{e}^*_{jt} \tilde{} N(0,1) \\ d_{jt} &= d_{jt-1} + n_{jt}, var(n_{jt}) = q_j \end{aligned}$$

We use the Particle Gibbs step described below to draw d_{jt} from its conditional posterior distribution. Given d_{jt}, the conditional posterior for q_j is inverse Gamma with degrees of freedom $T + vq_0$ and scale parameter $q_0 + n'_{jt}n_{jt}$.

A.3.2 Elements of $H_t\ H\left(H_t|C_t, Q_\Gamma, \Gamma_t, D_t, \breve{S}, \Phi_t, Q_\Phi, A_t, S, q\right)$

Given a draw for the time-varying parameters of the transition equation 17 and the time-varying parameters of the observation equation 12, the model has a multivariate non-linear state-space representation. Following recent developments in the seminal paper by Andrieu *et al.* (2010), we employ a particle Gibbs step to sample from the conditional posterior of $\tilde{h}_t$. Andrieu *et al.* (2010) show how a version of the particle filter, conditioned on a fixed trajectory for one of the particles can be used to produce draws that result in a Markov Kernel with a target distribution that is invariant. However, the usual problem of path degeneracy in the particle filter can result in poor mixing in the original version of particle Gibbs. Recent developments, however, suggest that small modifications of this algorithm can largely alleviate this problem. In particular, Lindsten *et al.* (2014) propose the addition of a step that involves sampling the 'ancestors' or indices associated with the particle that is being conditioned on. They show that this results in a substantial improvement in the mixing of the algorithm even with a few particles.[4]As explained in Lindsten *et al.* (2014), ancestor sampling breaks the reference path into pieces and this causes the particle system to collapse towards something different than the reference path. In the absence of this step, the particle system tends to collapse to the conditioning path. We employ particle Gibbs with ancestor sampling in this step.

Let $\tilde{h}_t^{(i-1)}$ denote the fixed the fixed trajectory, for $t = 1, 2, ..T$ obtained in the previous draw of the Gibbs algorithm. We denote the parameters of the model by Ξ, and $j = 1, 2, ..M$ represents the particles. The conditional particle filter with ancestor sampling proceeds in the following steps:

[4]See Nonejad (2015) for a recent application of this algorithm.

1. For $t = 1$
   (a) Draw $\tilde{h}_1^{(j)} \backslash \tilde{h}_0^{(j)}$,$\Xi$ for $j = 1, 2, ..M-1$. Fix $\tilde{h}_1^{(M)} = \tilde{h}_1^{(i-1)}$
   (b) Compute the normalised weights $p_1^{(j)} = \frac{w_1^{(j)}}{\sum_{j=1}^{M} w_1^{(j)}}$ where $w_1^{(j)}$ denotes the conditional likelihood: $\left|\Omega_1^{(j)}\right|^{-0.5} - 0.5 \exp\left(\tilde{e}_1 \left(\Omega_1^{(j)}\right)^{-1} \tilde{e}_1'\right)$ where $\tilde{e}_1 = Z_t - \left(c_t + \sum_{j=1}^{P} \beta_{t,j} Z_{t-j}\right)$ and $\Omega_1^{(j)} = A_t^{-1} H_1^{(j)} A_t^{-1'}$ with $H_1^{(j)} = diag\left(\exp\left(\tilde{h}_1^{(j)}\right)\right)$.
2. For $t = 2$ to T
   (a) Resample $\tilde{h}_{t-1}^{(j)}$ for $j = 1, 2, ..M-1$ using indices $a_t^{(j)}$ with $\Pr\left(a_t^{(j)} = j\right) \propto p_{t-1}^{(j)}$
   (b) Draw $\tilde{h}_t^{(j)} \backslash \tilde{h}_{t-1}^{(a_t^{(j)})}$,$\Xi$ for $j = 1, 2, ..M-1$ using the transition equation of the model (equation 17). Note that $\tilde{h}_{t-1}^{(a_t^{(j)})}$ denotes the resampled particles in step (a) above.
   (c) Fix $\tilde{h}_t^{(M)} = \tilde{h}_t^{(i-1)}$
   (d) Sample $a_t^{(M)}$ with $\Pr\left(a_t^{(M)} = j\right) \propto p_{t-1}^{(j)} \Pr\left(\tilde{h}_t^{(i-1)} \backslash \tilde{h}_{t-1}^{(j)}, \alpha_t, \theta_t, Q_t\right)$ where the density $\Pr\left(\tilde{h}_t^{(i-1)} \backslash \tilde{h}_{t-1}^{(j)}, \alpha_t, \theta_t, Q_t\right)$ is computed as $|Q_t|^{-0.5} - 0.5 \exp\left(\tilde{\eta}_t^{(j)} \left(Q_t\right)^{-1} \tilde{\eta}_t^{(j)}\right)$ where $\tilde{\eta}_t = \tilde{h}_t^{(i-1)} - \left(\alpha_t + \theta_t \tilde{h}_{t-1}^{(j)}\right)$. This constitutes the ancestor sampling step. If $a_t^{(M)} = M$ then the algorithm collapses to the simple particle Gibbs.
   (e) Update the weights $p_t^{(j)} = \frac{w_t^{(j)}}{\sum_{j=1}^{M} w_t^{(j)}}$ where $w_1^{(j)}$ denotes the conditional likelihood: $\left|\Omega_t^{(j)}\right|^{-0.5} - 0.5 \exp\left(\tilde{e}_t \left(\Omega_t^{(j)}\right)^{-1} \tilde{e}_t'\right)$ where $\tilde{e}_t = Z_t - \left(c_t + \sum_{j=1}^{P} \beta_{t,j} Z_{t-j}\right)$ and $\Omega_t^{(j)} = A_t^{-1} H_1^{(j)} A_t^{-1'}$ with $H_t^{(j)} = diag\left(\exp\left(\tilde{h}_t^{(j)}\right)\right)$.
3. End
4. Sample $\tilde{h}_t^{(i)}$ with $\Pr\left(\tilde{h}_t^{(i)} = \tilde{h}_t^{(j)}\right) \propto p_T^{(j)}$ to obtain a draw from the conditional posterior distribution

We use $M = 50$ particles in our application. The initial values μ_0 defined above are used to initialise step 1 of the filter.

B Appendix B: Calculating the marginal likelihood

The marginal likelihood is estimated via the Chib (1995) estimator. This estimator is based on rearranging the Bayes theorem. Consider the Bayes Theorem in logs:

$$H\left(\Psi, \Xi \backslash Z\right) = F\left(Z \backslash \Psi, \Xi\right) + P\left(\Psi\right) - H\left(Z\right)$$

where $\Psi = \{Q_\Gamma, \breve{S}, Q_\Phi, S, q\}$ i.e. parameters, Ξ denotes the state variables. $H\left(\Psi, \Xi \backslash Z\right)$ is the log posterior that equals the log likelihood $F\left(Z \backslash \Psi, \Xi\right)$ plus the log prior $P\left(\Psi\right)$ minus the log marginal likelihood $H\left(Z\right)$. Thus:

$$H\left(Z\right) = F\left(Z \backslash \Psi, \Xi\right) + P\left(\Psi\right) - H\left(\Psi, \Xi \backslash Z\right) \tag{23}$$

Because the parameters and the states do not appear in $H(Z)$, this expression can be evaluated at any point in the posterior density for Ψ. Conventionally this point is chosen to be the posterior mean Ψ^*.

Thus, to estimate the marginal likelihood we need to evaluate eq 23. $P(\Psi)$ is easiest to evaluate as it involves evaluating the prior for each parameter (at Ψ^*) and then taking the product. The likelihood $F(Z\backslash\Psi,\Xi)$ is evaluated using a Particle filter (see Arulampalam *et al.* (2002)). The key task is evaluating the posterior density $H(\Psi,\Xi\backslash Z)$. We turn to this next.

B.1 Evaluating the posterior density

The posterior density evaluated at the mean is defined as

$$\begin{aligned} H\left(Q_\Gamma^*,\breve{S}^*,Q_\Phi^*,S^*,q^*,\Xi\right) &= H\left(Q_\Gamma^*\backslash\breve{S}^*,Q_\Phi^*,S^*,q^*,\Xi\right)\times H\left(\breve{S}^*\backslash Q_\Phi^*,S^*,q^*,\Xi\right)\times \\ &\quad H\left(Q_\Phi^*\backslash S^*,q^*,\Xi\right)\times H\left(S^*\backslash q^*,\Xi\right)\times H\left(q^*\backslash\Xi\right) \end{aligned} \tag{24}$$

where $*$ denotes posterior mean. We suppress dependence on the data for notational simplicity. Each term on the RHS of equation 24 can be evaluated using an extended Gibbs sampler:

1. $H\left(Q_\Gamma^*\backslash\breve{S}^*,Q_\Phi^*,S^*,q^*,\Xi\right)$ is a complete conditional evaluated at the posterior mean for $\breve{S}^*,Q_\Phi^*,S^*,q^*$. The additional Gibbs run samples from 1) $H\left(\tilde{Q}_\Gamma\backslash\breve{S}^*,Q_\Phi^*,S^*,q^*,\tilde{\Xi}\right)$ and 2) $H\left(\tilde{\Xi}\backslash\tilde{Q}_\Gamma,\breve{S}^*,Q_\Phi^*,S^*,q^*\right)$ where the superscript ~denotes Gibbs draws. After a burn-in period, $H\left(Q_\Gamma^*\backslash\breve{S}^*,Q_\Phi^*,S^*,q^*,\Xi\right)\approx \frac{1}{M}\sum_{j=1}^{M}H\left(Q_\Gamma^*\backslash\breve{S}^*,Q_\Phi^*,S^*,q^*,\tilde{\Xi}\right)$ where M denotes the retained draws. Note that $H\left(Q_\Gamma^*\backslash\breve{S}^*,Q_\Phi^*,S^*,q^*,\tilde{\Xi}\right)$ is the inverse Wishart density with scale parameter calculated using this extended sampler.

2. $H\left(\breve{S}^*\backslash Q_\Phi^*,S^*,q^*,\Xi\right)$ can be written as

$$\int H\left(\breve{S}^*\backslash Q_\Phi^*,S^*,q^*,Q_\Gamma^*,\Xi\right)\times H\left(Q_\Gamma^*\backslash Q_\Phi^*,S^*,q^*,\Xi\right)dQ_\Gamma^*$$

   Thus this density can be approximated by a Gibbs sampler that samples from 1) $H\left(\tilde{\breve{S}}\backslash Q_\Phi^*,S^*,q^*,\tilde{Q}_\Gamma,\tilde{\Xi}\right)$, 2) $H\left(\tilde{Q}_\Gamma\backslash Q_\Phi^*,S^*,q^*,\tilde{\breve{S}},\tilde{\Xi}\right)$ and 3) $H\left(\tilde{\Xi}\backslash\tilde{Q}_\Gamma,Q_\Phi^*,S^*,q^*,\tilde{\breve{S}}\right)$. After a burn in period, the (inverse Wishart) density can be approximated as $\frac{1}{M}\sum_{j=1}^{M}H\left(\breve{S}^*\backslash Q_\Phi^*,S^*,q^*,\tilde{Q}_\Gamma,\tilde{\Xi}\right)$

3. $H\left(Q_\Phi^*\backslash S^*,q^*,\Xi\right)$ can be written as

$$\int\left[\int\underbrace{H\left(Q_\Phi^*\backslash Q_\Gamma^*,\breve{S}^*,S^*,q^*,\Xi\right)\times H\left(Q_\Gamma^*\backslash\breve{S}^*,S^*,q^*,\Xi\right)dQ_\Gamma^*}_{H\left(Q_\Phi^*\backslash\breve{S}^*,S^*,q^*,\Xi\right)}\right]\times H\left(\breve{S}^*\backslash S^*,q^*,\Xi\right)d\breve{S}^*$$

   This can be approximated by a Gibbs run that samples from (1) $H\left(\tilde{\breve{S}}\backslash\tilde{Q}_\Phi,S^*,q^*,\tilde{Q}_\Gamma,\tilde{\Xi}\right)$

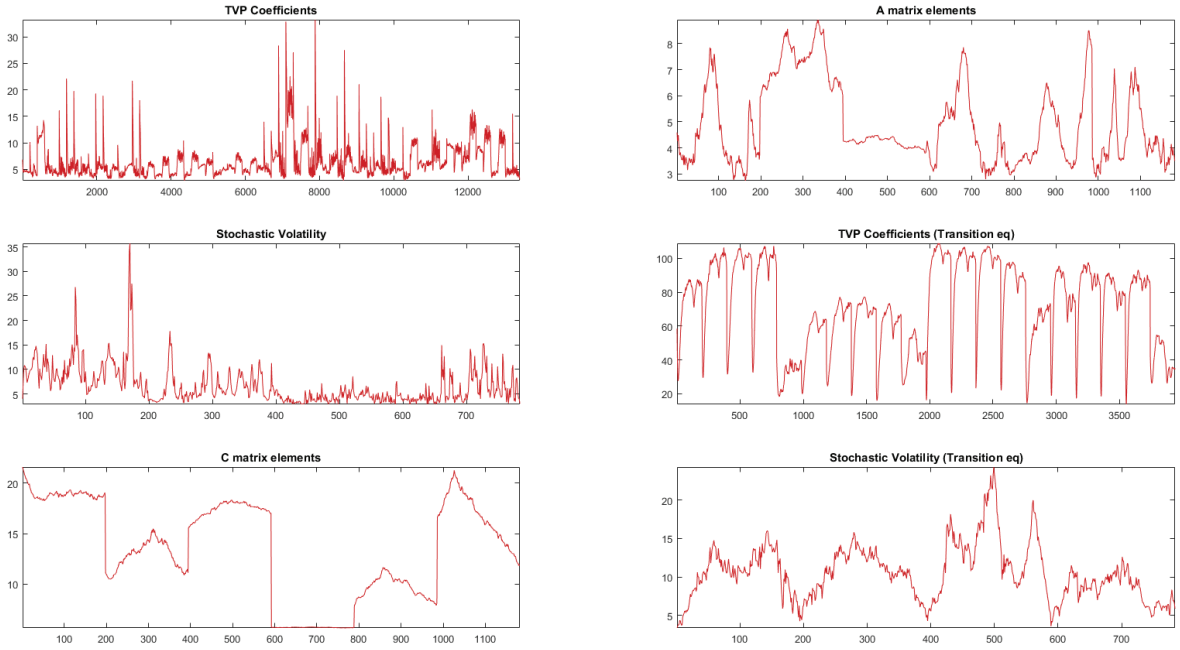

Figure 8: Inefficiency factors

(2) $H\left(\tilde{Q}_\Gamma \backslash \tilde{Q}_\Phi, S^*, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$ (3) $H\left(\tilde{\Xi} \backslash \tilde{Q}_\Gamma, \tilde{Q}_\Phi, S^*, q^*, \tilde{\breve{S}}\right)$ and (4) $H\left(\tilde{Q}_\Phi \backslash \tilde{Q}_\Gamma, S^*, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$.

After burn in we use the average $\frac{1}{M}\sum_{j=1}^{M} H\left(Q_\Phi^* \backslash \tilde{Q}_\Gamma, S^*, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$ as an approximation of the desired density. $H\left(Q_\Phi^* \backslash \tilde{Q}_\Gamma, S^*, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$ is inverse Wishart.

4. $H\left(S^* \backslash q^*, \Xi\right)$ can be approximated by a Gibbs sampler that samples from (1) $H\left(\tilde{\breve{S}} \backslash \tilde{Q}_\Phi, \tilde{S}, q^*, \tilde{Q}_\Gamma, \tilde{\Xi}\right)$ (2) $H\left(\tilde{Q}_\Gamma \backslash \tilde{Q}_\Phi, \tilde{S}, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$ (3) $H\left(\tilde{\Xi} \backslash \tilde{Q}_\Gamma, \tilde{Q}_\Phi, \tilde{S}, q^*, \tilde{\breve{S}}\right)$ (4) $H\left(\tilde{Q}_\Phi \backslash \tilde{Q}_\Gamma, \tilde{S}, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$ (5) $H\left(\tilde{S} \backslash \tilde{Q}_\Phi, \tilde{Q}_\Gamma, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$. After a burn-in period we use the average $\frac{1}{M}\sum_{j=1}^{M} H\left(S^* \backslash \tilde{Q}_\Phi, \tilde{Q}_\Gamma, q^*, \tilde{\breve{S}}, \tilde{\Xi}\right)$

5. $H\left(q^* \backslash \Xi\right)$ can be approximated as $\frac{1}{M}\sum_{j=1}^{M} H\left(q^* \backslash Q_\Gamma, \breve{S}, Q_\Phi, S, \Xi\right)$ using draws from the original Gibbs sampler

C Appendix C: Convergence

Figure 8 reports inefficiency factors for key parameters calculated using a Parzen window. For most parameters, the estimates are close to or below the recommended level of 20, with the time-varying

coefficients of the transition equation providing some exceptions. Given the heavily parameterised nature of the model, these results provide reasonable evidence for convergence of the algorithm.

School of Economics and Finance

This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/