The Impact of Dynamic Covariance on Impulse Response Functions: Applications to Structural VARs and DSGE Models

Abstract

This paper presents a unified framework for the computation of impulse response functions for Vector Autoregressive (VAR) models with multivariate generalized autoregressive conditionally heteroskedastic disturbances (MV-GARCH). The solutions obtained from solving the quadratic form as given by the second moment equation should be used to adjust the impulse response functions (IRFs) from the mean equations. The adjusted IRFs often result in significantly different time profiles from the unadjusted impulse responses. This approach provides for a system consistent solution for multivariate linear autoregressive models, with time varying second moments.

Keywords: Macroeconomic Time Series Modelling, Multivariate GARCH, Vector Autoregression, Impulse Response
JEL Classification: G50, C15, C22

Vector Autoregressive (VAR) models in various forms are one of the most extensively used statistical procedures for both forecasting and policy evaluation. Their use is somewhat controversial as their economic specification is deemed too general for the testing of exact parametric restrictions derived from fully developed economic models based on the existence of the representative agent. Despite such shortcomings VAR models can incorporate restrictions that help to identify clearly the causal chains between variables and the nature of innovations. By and large such models have gained acceptance in the light of their comparative performance given the relative simplicity of the implementation and forecasting ability. However, the recent economic turbulence since 2007 casts serious doubts on the ability of all macroeconomic models to account for the response of the economy to the policy impulses during the crisis. We take it as indicative that such shortcomings, in the context of VAR models, stem from the omission of relevant information affecting their specification. Such omissions subsequently affect the resulting dynamic adjustments as measured by the impulse response functions (IRFs).

This paper offers some useful general results when the disturbance term in a VAR model exhibits time series dependency in variance, specifically the errors are multivariate generalized autoregressive conditionally heteroskedastic (MV-GARCH). We show that there are substantial differences in the impulse response functions between VAR models, in which such effects are suppressed and those that are fully specified. We estimate three well established macroeconomic models, first the interaction between inflation and GDP for the US reported by Bullard and Keating (1995) our second the structural VAR model of Blanchard and Quah (1989) and compare the resulting IRFs with and without the incorporation of MV-GARCH effects for post war data and finally we apply the compute the IRFs for Dynamic Stochastic General Equilibrium (DSGE) model with four state variables which is expressed in terms of structural VAR(1),under two alternative multivariate-covariance structures,

This method is valuable to academics and policy makers utilizing vector autoregressive models containing data with possible multivariate autoregressive conditionally heteroscedastic disturbances who wish to ascertain the response of the system to innovations. Consider the following scenario, a simple stationary bivariate vector autoregression with an impulse response decay of 50% of one standard deviation when simulating a response in the first equation to a unit shock in equation two and a 25% persistence in its autoregression. The response in variable one is measured in numbers of long run standard deviations and we would expect to see a smooth return to equilibrium after the shock and we could plot the impulse response function (IRF) using the multitude of methods suggested in the literature.

However, all of these impulse responses will be under the assumption that the expected standard deviation measure at $t + s$, where s is the number of steps is the same as at t when the impulse was initiated. If the model exhibits MV-GARCH effects then the shock will generate a response in the MV-GARCH model that will change the subsequent measure of the standard deviations. If we assume that the long run standard deviation is 30% and the MV-GARCH effect raises it to 50% at $t + 1$, then the shock, measured in terms of t standard deviations will not have dissipated as fast as would have been expected under a time invariant variance/covariance structure.

Furthermore, for models with structural restrictions the adjustment to the variance-covariance matrix has a profound impact on the derivation of the structural parameters and hence the latent variables implied from the structural framework. Our contribution is in postulating a new impulse response function that factors in these adjustments and is easily incorporated into an SVAR framework. Initially we outline two uses of this impulse response function, one in a basic setting with no structural restrictions and second with a simply time invariant structural restriction imposed. We extent the analysis to a structural

VAR(1) derived from a well known , calibrated DSGE , where we compute , using calibrated parameters, the impact of shocks to the state variables with and without MV-GARCH effects describing the system's dynamic covariance.

Univariate and multivariate variance persistence is observed in a great number of processes, including the well documented examples of inflation and interest rates. There is prima facie evidence to suggest that volatility clustering may even be present in output, see for instance DeJong et al. (2006), Lee (2006) and Engle and Rangel (2008). As each of these variables are fundamental to macro policy evaluation the correct representation of their response to a shock is of great importance.

The response of non-linear econometric structures to shocks has been addressed in a number of papers, most notably Gallant et al. (1993) and Koop et al. (1996). The former develops the concept and derives rigorously what is termed as *generalized non-linear impulse response functions*. It is suggested that such a concept can deal adequately with a series of problems that arise when models depart from the calculation of the responses within the linear autoregressive paradigm. Gourieroux and Jasiak (1999), extract the nonlinear residual structures as proxies for the non-linear innovations and develop a set of test statistics that can be used in testing the specification of the underlying shock response structure.

Gallant et al. (1993) provide a methodology for computing the impulse response for a non-linear time series model by computing the differences between *the baseline approach* and the conditional moment profile, using a semi-parametric methodology. The return volatility-trading volume relationship, which is postulated in terms of conditional second moments in the returns and conditional first moments in volume, is used as an application.

The growing acceptance of multivariate-generalized autoregressive conditionally heteroscedastic, henceforth MV-GARCH, models has led to the requirement for calculating the impulse response of the conditional variance after a shock. Lin (1997) derives such a measure in the context of a MV-GARCH model and evaluates the small sample properties of the 95% standard errors that surround the IRF point estimates by means of a Monte Carlo study utilizing a constant correlation type MV-GARCH(1,1).

Impulse response functions and MV-GARCH models have a limited history, the VAR-MV-GARCH models of Sin (2005), Polasek and Ren (1998) and Polasek and Ren (2000) illustrate a mechanism for computing an impulse response function in variance for a general VAR model. Elder (2003) offers a description of an algorithm for defining impulse responses in variance. Hafner and Herwartz (2001) demonstrate a model of impulse response in variance and apply it to exchange rates, utilizing the impulse responses as a guide to system identification. However none of the current derivations of impulse responses in variance incorporate the dynamics of a fully specified MV-GARCH model. When computing the IRF in the context of a VAR-MV-GARCH exclusively from the mean equation, the information from the second moments is neglected as it is implicitly assumed that the system covariance matrix is fixed in the presence of a shock or innovation. However, the very presence of MV-GARCH is indicative that the time invariance assumption regarding the covariance matrix cannot be sustained.

To incorporate the fact that the conditional covariance matrix is evolving through time (deterministically) the appropriate manner to compute the IRF is as follows: for any given shock, compute the expectation of the second moments (the quadratic form) of the system and compare it to the quadratic form that would have existed without the shock. In the context of a VAR-MV-GARCH model the shock will feed into the mean equations via the triangular decomposition of the covariance matrix that is now time dependent. This provides corrected *in-mean* IRFs that are computed from an evolving covariance matrix rather than a constant one as is the case when the system ignores any MV-GARCH effects.

Jorda (2005) simulates a multi equation model with a single GARCH model in-mean and in-variance, using an innovative local projections approach. The impulse responses are calculated by a sequence of local projections that are robust to the underlying linear or non-linear structure of the data generation process and this allows for statistical inference of the directly computed impulse response coefficients from the regression error structure.

In terms of motivation for the inclusion of time varying second moments in macro models, recent work by Bloom (2009) has demonstrated, by use of stylised simulations, that uncertainty shocks play very important roles in investment decision making and that there is an interaction between direction and uncertainty in aggregate models of output. We show how innovations in the mean system equations affect the variance/covariance system and such effects should be included in conjunction with the underlying model in mean. We demonstrate this by solving a matrix polynomial function and show how the interactions between the first and second moments in a dynamic model can be expressed in a mutually consistent manner.

The paper is organized as follows: §(1) and §(2) provide the algorithm for the computation of impulse response functions for VAR-MV-GARCH models and propose an inferential procedure for computing error bounds around the point estimate. §(3) outlines two estimated macroeconomic examples, illustrating how this methodology can be used to augment existing research. §(4) We use a calibrated DSGE model and compare the IRFs without and in the presence of MVGARCH(1,1) effects in two of the four state variables. Our concluding remarks and suggestions for uses of this approach are provided in §(5).

1. VAR Models with MV-GARCH Disturbances

Consider the general unrestricted stationary r-order linear VAR model of a covariance stationary process $y_t \in \mathbb{R}^n$. We will initially assume a general autoregressive framework where the conditional first and second moments are generated by a pair of autoregressive processes with parameters θ,

$$C*(L)\,y_t = w\Sigma_t + c + C(L)\,Q_t\varepsilon_t \quad (1)$$

$$S(L)\,\Sigma_t = QQ' + G(L)\,A_t \quad (2)$$

where $Q_t\varepsilon_t\varepsilon_t'Q_t' = A_t$ and $Q_tQ_t' = \Sigma_t$, for a matrix Q_t. For estimation purposes it is useful to restrict Q_t to being upper triangular. However, for the purposes of computing the impulse response functions we are agnostic to this imposition, simply requiring that the observed deviations from the conditional predictor are specified by $y_t - \mathbb{E}_t\left(y_{t|t-1,\ldots,t_0}\right) = Q_t\varepsilon_t$. In A we outline two common MV-GARCH specifications and provide guidance on parameter configurations that result in impulse responses with finite deviation from equilibrium.

The vector c and matrix QQ' are constants, in the first and second moment equations. The elements of the white noise vector ε_t are denoted $\varepsilon_{i,t}$. For estimation purposes they are assumed to be a standard Markov process with the following moments, $\mathbb{E}(\varepsilon_{i,t}) = 0$, $\mathbb{E}\left(\varepsilon_{i,t}^2\right) = 1$, $\mathbb{E}(\varepsilon_{i,t}\varepsilon_{j,t}) = 0, \forall j \neq i$ and $\mathbb{E}(\varepsilon_{i,t}\varepsilon_{j,t-k}) = 0, \forall j, k = -\infty, \ldots, \infty$.

For simplicity we assume that the vector autoregressive moving average representation in the first moment is linear, i.e. $C*(L) = I - \sum_{i=1}^{r} \Pi_i L^i$ and $C(L) = I - \sum_{i=1}^{s} \Pi_{0,i} L^i$. The variance-covariance generating process may be more general, i.e. $S(L) = 1 - \sum_{i=1}^{p} f^i L^i$ and $G(L) = 1 - \sum_{i=1}^{q} g^i L^i$, where $f^i : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ and $g^i : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ are arbitrary matrix functions. The vector of weights w impose a functional link between the deterministic terms in the first and second moment equations. For most of our applications we ignore the finite order moving average terms and restrict $w = 0$.

For tractability of computing the impulse response functions for our examples we will require that $A^{-1}(L)\,C(L)$ exists (i.e. we restrict ourselves to stationary vector processes). The vector autoregression is therefore invertible. We now introduce two final notational representations of the vector process. First, given the invertibility condition,there is the infinite moving average representation denoted $C^{\infty}(L)\,y_t - \tilde{c} = \tilde{\varepsilon}_t$. Second, using the standard approach we rewrite the r order autoregressive process as a first order VAR model of the form $\tilde{v}_t(1 - \tilde{\Phi}L) = \tilde{\epsilon}_t$, where $v_t = [y_t', y_{t-1}', \ldots, y_{t-(r-1)}']'$ and $\tilde{\epsilon}_t = [\epsilon_t', 0'_{r-1}]'$, where 0_{r-1} is an $r-1$ length vector of zeros.

By construction, if $\Pi = [\Pi_1', \Pi_2', \ldots, \Pi_r']'$ then the square (and transposed if we consider the tradational presentation) companion matrix Φ is defined as follows,

$$\underset{nr\times nr}{\Phi} = \begin{bmatrix} \underset{n\times n(r-1)}{\Pi'} \\ \underset{n(r-1)\times n(r-1)}{I}, \underset{n(r-1)\times n}{0} \end{bmatrix} \tag{3}$$

Where I and 0 are, respectively an appropriately sized identity matrix and an appropriately sized matrix of zeros. Stationarity in the mean component implies that the eigenvalues of Φ must lie within the unit circle for an infinite moving average representation $C^{\infty}(L)\,v_t = \tilde{\epsilon}_t$ to exist.

The standard impulse response functions for an unrestricted VAR are computed from the innovations of the polynomial powers of this matrix, denoted $A^{(s)} = A \times A \ldots_k \times A$, where A is a square matrix and the integer s is the desired number of innovations, therefore if $F_s = \Phi^{(s)}$ and $F_{11,s}$ is the $n \times n$ matrix of entries in the upper left of F_s then the Wald innovations of the unrestricted VAR will be the entries of this matrix with the first row representing the innovations for the first equation, the second for the second equation and so on.

If the first moment is stationary, then the process generating the second moment, must also be stationary, therefore $\Sigma = \mathbb{E}(\Sigma_t) = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\Sigma_t$, subsequently the long run covariance matrix is factorised as $\Sigma = \tilde{Q}\tilde{Q}'$.

Thus far we have not made any distributional assumptions regarding vector of innovations ε_t, assuming that $\varepsilon_{i,t} \sim \mathcal{N}(0,1)$, then the time t log likelihood for a candidate θ, denoted $\tilde{\theta}$ will be

$$\mathscr{L}_t[\tilde{\theta}] = -\tfrac{1}{2}\left(n\log(2\pi) + \left|\tilde{\Sigma}_t;\tilde{\theta}\right| + \left(y_t - \tilde{y}_t;\tilde{\theta}\right)\tilde{\Sigma}_t^{-1};\tilde{\theta}\left(y_t - \tilde{y}_t;\tilde{\theta}\right)\right) \tag{4}$$

for a set of parameters θ describing the system of equations. Setting, $\tilde{y}_t$ and $\tilde{\Sigma}$ to be the conditional predictors of y_t and Σ_t for a given set of parameters $\tilde{\theta}$, the log likelihood for a sample of length T is therefore, $\mathscr{L}[\tilde{\theta}] = \sum_{t=1}^{T} \mathscr{L}_t[\tilde{\theta}]$.

Estimation proceeds via $\theta^* = \arg\min_{\tilde{\theta}} -\mathscr{L}[\tilde{\theta}]$. Interestingly, most software implementations of this type of model use numerical gradients and numerically evaluated Hessian matrices, however in certain cases analytic first and second derivatives are available (or can be easily approximated by a polynomial expansion), the degree of improvement in attaining the minima and the accuracy of computing standard errors is substantial.

We now have two difference cases, first where $w \neq 0$ and second $w = 0$. Setting $\theta = [\theta'^{y}, \theta'^{\Sigma}]'$, the gradient and Hessian of the log likelihood function are noted as $\nabla\mathscr{L}[\tilde{\theta}]$ and $\mathscr{H}^{-1}[\theta^*]$. Setting $\bar{\theta}$ as the 'true' parameters for the model, the error variance-covariance matrix of the estimated parameters is $\mathrm{cov}(\theta^* - \bar{\theta}) = T^{-1}\mathscr{H}^{-1}[\theta^*]$.

For the general case where w is unrestricted, the parameters for the mean model are divided as follows $\theta^{y} = [\theta'^{y,y}, \theta'^{,y,\Sigma}]'$, the Hessian is therefore:

$$\mathscr{H}[\theta^*] = \begin{bmatrix} \frac{\delta^2\mathscr{L}[\tilde{\theta}]}{\delta\theta^{y,y}\delta\theta'^{y,y}} & \frac{\delta^2\mathscr{L}[\tilde{\theta}]}{\delta\theta^{y,\Sigma}\delta\theta'^{y,y}} & 0 \\ \frac{\delta^2\mathscr{L}[\tilde{\theta}]}{\delta\theta^{y,\Sigma}\delta\theta'^{y,y}} & \frac{\delta^2\mathscr{L}[\tilde{\theta}]}{\delta\theta^{y,\Sigma}\delta\theta'^{y,\Sigma}} & 0 \\ 0 & 0 & \frac{\delta^2\mathscr{L}[\tilde{\theta}]}{\delta\theta^{y,y}\delta\theta'^{y,y}} \end{bmatrix} \tag{5}$$

when $w = 0$, there are no parameters linking the mean and variance equations directly and as such the hessian simplifies to

$$\mathscr{H}[\theta^*] = \begin{bmatrix} \frac{\delta^2\mathscr{L}[\tilde{\theta}]}{\delta\theta^{y,y}\delta\theta'^{y,y}} & 0 \\ 0 & \frac{\delta^2\mathscr{L}[\tilde{\theta}]}{\delta\theta^{\Sigma,\Sigma}\delta\theta'^{,\Sigma,\Sigma}} \end{bmatrix} \tag{6}$$

whilst the imposition of $w = 0$, results in the separation of the first and second equation for estimation purposes, the impact of shocks to $\varepsilon_{i,t}$ are transmitted through adjustments in the volatility matrix Q_t.

2. Impulse Response Functions

Our main object of interest is to evaluate the matrix of partial derivative $\partial y_{t+s}/\partial\varepsilon'_t$. To accomplish this we need to also compute the following matrix and array of partial derivatives, $\partial y_{t+s}/\partial Q_t\varepsilon_t$ and $\partial Q_{t+s}/\partial\varepsilon'_t$. For the purposes of this section we shall assume that we are interested in an impulse that originates at time t in ε. For brevity we shall interchange between shock and innovation, both of which in our context refer to a one unit impulse to an element in ε_t.

We now impose the following notational forms for the impulse responses. Let $\tilde{\Xi}_s = [\partial y_{t+s}/\partial \varepsilon_{i,t}]_{i=1}^{n}$, to be the matrix of responses at $t+s$ to an innovation at time t in the vector ε_t. For the observed shocks $Q_t \varepsilon_t$, we set $\tilde{q}_{i,t}$ to be the i^{th} row of Q_t therefore the matrix of impulse responses for the observed shocks is $\tilde{\Psi}_s = [\partial y_{t+s}/\partial \tilde{q}'_{i,t}\varepsilon_t]_{i=1}^{n}$. Finally for each element of ε_t, there is a matrix of responses in Q_t contained in the array $\tilde{\mathcal{S}}_s = [\partial Q_{t+s}/\varepsilon_{i,t}]_{i=1}^{n}$.

Computing the sequence of Q_{t+s}, in response to a shock in ε_t is the simplest part of the computation. First, we must select the system's steady state variance. Unfortunately, we cannot assume that $L^i\Sigma_t = 0, \forall i \in 1, \ldots, p$ as the past covariance matrices are by definition non-zero. A natural approach is to choose $L^i\Sigma_t = \Sigma$, i.e. the long run covariance matrix, however, this would require assumptions regarding the the terms included in $G(L)A_t$. We choose the following, first we assume that $\Sigma_t = \Sigma$. Second we assume that $L^iA_t = 0, \forall i = 1, \ldots, p$. Finally, we specify that the set of p matrices $\{L^i\Sigma_t\}_{i=1}^{p}$ is at a minimum distance (in the Euclidean sense of the Frobenius norm of the difference) to Σ subject to the prior constraints. Formally,

$$\left\{L^i\Sigma_t^*\right\}_{i=1}^{p} = \underset{L^i\Sigma_t}{\arg\min}\left[\sum_{i=1}^{p}\left|L^i\Sigma_t - \Sigma\right|_F\right] \tag{7}$$

$$\text{subject to} \quad \Sigma - QQ = \sum\nolimits_{i=1}^{p} f^i L^i\Sigma_t \tag{8}$$

$$\alpha_i\left(L^i\Sigma_t\right) = \Sigma \tag{9}$$

where $\alpha_i(\cdot)$ are affine functions such that Equation 7 has a unique solution. The nature of $\alpha_i(\cdot)$ is dependent on the functional form of f^i. Typically, for simple matrix processes $\alpha_i\left(L^i\Sigma_t\right) \equiv \tilde{\alpha}_i L^i\Sigma_t$ provides enough restrictions to identify exactly $L^i\Sigma_t^*$.

For most choices of MV-GARCH model the steady state covariance does not affect $\partial Q_{t+s}/\partial\varepsilon'_t$. However, the starting covariance matrix Σ_t does influence the magnitude of the response in $\partial y_{t+s}/\partial Q_t\varepsilon_t$ and hence $\partial y_{t+s}/\varepsilon_t$.

We now specify two states; first the steady state suggested above, i.e. $\Sigma_t^0 = \Sigma | A_t^0, A_{t-1}^0, \ldots, A_{t_0}^0$, second the state containing the innovation is ε_t denoted as $\varepsilon_t = \delta$. Subsequently we define $\Sigma_t^\delta | A_t^\delta, where A_{t-1}^0, \ldots, A_{t_0}^0$ and $A_t^\delta = \tilde{Q}\varepsilon_t^{\delta,\prime}\varepsilon_t^\delta\tilde{Q}'$ and $A_{t-i}^0 = Q_i^*\varepsilon_{t-i}^0\varepsilon_{t-i}^{\prime 0}Q_i^{\prime *}$, where $Q_i^* = \tilde{Q}$ when $i = 0$ and $Q_i^*Q_i^{\prime *} = L^i\Sigma_t^*$ for $i \in 1, \ldots, p$. The innovations are defined as follows; for a given shock at time t, $\varepsilon_t^\delta = \delta$, we denote zero innovations by $\varepsilon_t^\delta = 0$. Typically, to identify reactions to an innovation in a single equation we set one element of δ to unity, with the rest equal to zero. However for covariance shocks, we need to set pairs of elements to unity.

The initial response in the covariance matrix is therefore measured by $\Delta_0 = \Sigma_t^\delta - \Sigma_t^0$. Similarly we set $Q_t^\delta Q_t'^\delta = \Sigma_t^\delta$ and $Q_t^0 Q'^0_t = \Sigma_t^0$. Setting $\tilde{q}_{i,j,t}$ to be an element of Q_t, then $\left(\tilde{q}_{i,j,t}^\delta - \tilde{q}_{i,j,t}^0\right)\delta^{-1} = \partial\tilde{q}_{i,j,t}/\partial\varepsilon_t$ is the initial direction in Q_t in response to an innovation in ϵ_t.

Similarly we can now project Q_t forwards s steps, whereby $(\tilde{q}_{i,j,t+s}^\delta - \tilde{q}_{i,j,t+s}^0)\delta^{-1} = \partial\tilde{q}_{i,j,t+s}/\partial\varepsilon_t$. At each forward step the conditional covariance matrix changes, therefore, given the stationarity condition in covariance, $\lim_{s\to\infty}\Delta_s = 0$. We can now set out a systematic construction of the array $\tilde{\mathcal{S}}_s$. First we can index $\delta_{\mathbf{i}}$ where $\mathbf{i}$ denotes that each element of δ is equal to zero except element $i \in 1,\ldots,n$. For simplicity we will ignore covariance shocks, therefore

$$\tilde{\mathcal{S}}_s = \left[Q_{t+s}^{\delta_{\mathbf{i}}} - Q_{t+s}^0\right]_{i=1}^n . \tag{10}$$

Next, we compute the observed adjustment in y_t to a shock in $Q_t\varepsilon_t$. It is important to note that whilst the impulse responses collected in $\tilde{\mathcal{S}}_s$ are for most forms of f^i unaffected by the imposition of $\Sigma_t = \Sigma$, the responses in mean will be dependent on this assumption. Similarly to the previous case in variance we set $y_t^0 = y_t | Q_t\varepsilon_t^0,\, Q_{t-1}\varepsilon_{t-1}^0, \ldots, Q_{t_0}\varepsilon_{t_0}^0$ and $y_t^\delta = y_t \left| Q_t\varepsilon_t^\delta, Q_{t-1}\varepsilon_{t-1}^0, \ldots, Q_{t_0}\varepsilon_{t_0}^0\right.$ and as we have imposed previously $Q_t = \tilde{Q}$. The identification of Q_{t-i} is not relevant given that $\epsilon_t^0 = 0$, however, for completeness we set $Q_{t-i} = Q_i^*$. Setting $\tilde{q}$ to be a row of $\tilde{Q}$, then the size of an impulse generated by $\delta_{\mathbf{i}}$ will be $\tilde{q}'\delta_{\mathbf{i}}$.

Using the established notation for δ the matrix of impulse responses $\tilde{\Psi}_s$ measures in the units of y_t, will have rows indexed by i from innovations described in $\delta_{\mathbf{i}}$. Therefore

$$\tilde{\Psi}_s = \left[\left(y_{t+s}^\delta - y_{t+s}^0\right)\left(\tilde{q}_i'\delta_{\mathbf{i}}\right)^{-1}\right]_{i=1}^n \tag{11}$$

Which are the normal generalised impulse response functions computed from the long run covariance matrix Σ.

We can now rewrite the impulse in y_t as follows; consider a state predictor of $y_{t+s}^0 = y_{t+s}|\varepsilon_t = 0,\ldots,\varepsilon_{t+s} = 0$, this is the steady state predictor with no innovations. We then compute the quantities $y_{t+s}^\delta = y_{t+s}|Q_t\varepsilon_t^\delta, Q_{t+1}\ \varepsilon_{t+1}^0, \ldots, Q_{t+s}\varepsilon_{t+s}^0$ and $y_{t+s}^0 = y_{t+s}|Q_t\varepsilon_t^0$, $Q_{t+1}\varepsilon_{t+1}^0, \ldots, Q_{t+s}\varepsilon_{t+s}^0$, and subsequently the distance $(y_{t+s}^\delta - y_{t+s}^0)'(\tilde{q}_{i,t+s}^\delta)^{-1} = \tilde{u}_{i,t+s}$ when divided by the initial vector of innovations δ will be the impulse response in the vector y_t to a shock to ε_t. Therefore the vector of responses in y_t at s periods into the future to a shock to an element of ε_t, maybe written as $\partial y_{t+s}/\delta\varepsilon_{i,t} = \tilde{u}_{t+s}\delta_i^{-1}$. We now address the issue of these changes in the units of $\tilde{q}^{\partial}_{i,t+s}$.

The units of δ are in terms of the inverse of the factor Q_t. However, the $y^{\delta}_{t+s} - y^0_{t+s}$ is measured in the units of the original variables with deviations driven by $Q_t\delta$, entering into the mean equation. By constructing the Q^{δ}_{t+s} as an adjustment from Q^0_{t+s} the rows of Q^{δ}_{t+s}, $\tilde{q}^{\partial}_{i,t+s}$ will scale each adjustment in the vector y_{t+s} to the original innovations from δ.

This adjustment is important for the following reason, observed shocks to the vector process y_t, i.e. $y_t - \mathbb{E}_t\left(y_{t|t-1,\ldots,t_0}\right) = Q_t\varepsilon_t$ are translated from the white noise process ε_t via the factorised matrix Q_t. Therefore identification of ε_t is through the estimation of the latent matrix Q_t, as such $(y_t - \mathbb{E}_t(y_{t|t-1,\ldots,t_0}.))Q_t^{-1} = \varepsilon_t$.

Setting the initial innovation at t to $y^{\delta}_t - y^0_t = \tilde{u}_{i,t} = Q_t\delta$, the standardised deviations are therefore $\left(y^{\delta}_t - y^0_t\right)Q_t^{-1} = \delta$. The shock now enters the system of equations as such $\left(y^{\delta}_{t+s} - y^0_{t+s}\right)Q_t^{-1} = \delta_s$. We can interpret δ_s as the remaining component of the impulse ascribed directly to δ rather than to the subsequent realisation in covariance through $Q^{\delta}_{t-i}\delta$. As the innovations in $\delta_{s,\mathbf{i}}$ are unit in magnitude the matrix of standardised impulse responses is therefore:

$$\Xi_s = \left[\left(y^{\delta_{\mathbf{i}}}_{t+s} - y^0_{t+s}\right)\left(Q^{\delta_{\mathbf{i}}}_t\right)^{-1}\right]_{i=1}^{n} \tag{12}$$

$$= \left[\left(y^{\delta_{\mathbf{i}}}_{t+s} - y^0_{t+s}\right)\left(Q_t\right)^{-1}\right]_{i=1}^{n} \tag{13}$$

where $\delta_{\mathbf{i}}$ indexes each set of impulse response functions, for a given innovation at t.

However, this rescaled response, is measured in terms of the factorised covariance matrix $Q^{\delta_{\mathbf{i}}}_t = Q_t$ as the shock in ε takes one iteration to enter the covariance generating system. Reconstructing the adjustment in y_t at s-steps to a shock in ε_t, we rescale $\delta_{s,\mathbf{i}}$ to $\partial y^{\delta_{\mathbf{i}}}_{t+s} = (Q^{\delta_{\mathbf{i}}}_{t+s})^{-1}\delta_{s,\mathbf{i}}$, and the impulse responses are collected to give the following time varying matrix

$$\tilde{\Xi}_s = \left[Q^{\delta_{\mathbf{i}}}_{t+s}\left(y^{\delta_{\mathbf{i}}}_{t+s} - y^0_{t+s}\right)\left(Q_t\right)^{-1}\right]_{i=1}^{n}. \tag{14}$$

These are the structurally adjusted generalised impulse responses. They are measured in the original units of the elements of y_t, and are recomputed to account for the change in covariance matrix $\Sigma_t = Q_tQ_t'$. Even when the restriction $w = 0$ is imposed, as $\Sigma_{t+s} \neq \Sigma$, as

$\partial y_{t+s}^{\delta_\mathbf{i}}$ will differ from the case when Σ_{t+s} is assumed to equal Σ for all forward steps.

2.1. Standard Errors for Impulse Responses

The asymptotic impulse responses derived via the 'delta method' are not easy to replicate in this set-up, therefore we utilize a monte-carlo simulation approach sampling directly from the error covariance matrix of the simultaneously estimated parameters. The impulse response functions derived previously are dependent on the set of parameters θ that describe the time series properties of 1. These parameters are estimated from data with error and this needs to be incorporated in the impulse response functions. Consider the impulse responses at s collected in the matrix $\tilde{\Xi}_s$. We define $\tilde{\Xi}_s^*$ to be the standardised impulse responses at $t+s$ computed from the point estimate of θ^*. For a vector draw

$$\theta^k \sim \mathcal{N}\left(\theta^*, T^{-1}\mathscr{H}^{-1}[\theta^*]\right) \tag{15}$$

the corresponding impulse responses at $t+s$ will be $\tilde{\Xi}_s^i$ and $\tilde{\xi}_{i,j,s}^k$ is the i,j element of $\tilde{\Xi}_s$ for the k-draw of θ^k. The sample distribution of $\tilde{\xi}_{i,j,s}^k$ will not be normal therefore we compute standard error bounds using the empirical cumulative density function derived from sorting $\tilde{\xi}_{i,j,s}^k$. For our applications we plot the 95% mass from the lower 2.5 percentile to the upper 97.5 percentile.However, the asymptotic distribution can be shown to converge via the delta method as follows. From standard asymptotic theory we can write the convergence in distribution of the parameter error $\bar{\theta}-\hat{\theta}$ as follows

$$\sqrt{T}(\bar{\theta}-\hat{\theta}) \to^d \mathcal{N}\left(0, \mathscr{H}^{-1}[\theta^*]\right)$$

let $\mathscr{R}_{ijs} : \theta \in \Theta \to \mathbb{R}$ be a vector operator, such that $\mathscr{R}_{ijs}[\theta]$ is the s step of the ij impulse response from previous for a given parameter vector θ within the compact set of viable parameter vectors Θ. If we set $\nabla\mathscr{R}_{ijs}[\theta]$ to be the gradient vector of $\mathscr{R}_{ijs}$ with respect to theta then the asymptotic distribution of impulse response function ij at step s will be:

$$\sqrt{T}(\mathscr{R}_{ijs}[\bar{\theta}]-\mathscr{R}_{ijs}[\hat{\theta}]) \to^d \mathcal{N}\left(0, \nabla'\mathscr{R}_{ijs}[\theta]\mathscr{H}^{-1}[\theta^*]\nabla\mathscr{R}_{ijs}[\theta]\right)$$

therefore our sampling choice for the monte-carlo standard errors is asymptotically consistent. The major issue comes from the evaluation of $\nabla'\mathscr{R}_{ijs}[\theta]$ which needs to be evaluated numerically for most types of volatility model.

2.2. A Specific Example of the Algorithm

A natural question arises as to why this is a useful decomposition? For standard VAR models the new impulse response, simply captures a natural adjustment for the change in spread, subsequent to a shock in ε. We will now outline the 'base-case' impulse response function for the simplest VAR system a centered bivariate first order VAR model with a Cholesky decomposition of Σ_t. Let the evolution of y_t be given by the following pair of equations

$$y_{1,t} - \pi_{1,1}y_{1,t-1} - \pi_{1,2}y_{2,t-1} = q_{1,1,t}\varepsilon_{1,t} + q_{1,2,t}\varepsilon_{2,t} \quad (16)$$

$$y_{2,t} - \pi_{2,1}y_{1,t-1} - \pi_{2,2}y_{2,t-1} = q_{2,2,t}\varepsilon_{2,t}. \quad (17)$$

where $\pi_{1,1}, \pi_{1,2}, \pi_{2,1}, \pi_{2,2}$ are parameters and $0.5(\pm(\pi_{1,1}^2 - 2\pi_{1,1}\pi_{2,2} + 4\pi_{1,2}\pi_{2,1} + \pi_{2,2}^2)^{0.5} + \pi_{1,1} + \pi_{2,2})$ lies within the unit circle.

The importance of ordering is immediately apparent as shocks to $\varepsilon_{2,t}$ enter both the first and second equations in this set up, via $q_{1,2,t}$. The time evolution of Σ_t is assumed to be a first order BEKK model, see Engle and Kroner (1995). Therefore, the three elements of the covariance matrix are computed from the following matrix system,

$$q_{1,1,t}^2 + q_{1,2,t}^2 = q_{1,1}^2 + q_{1,2}^2 + a_1^2\left(q_{1,1,t-1}^2 + q_{1,2,t-1}^2\right) + b_1^2\left(q_{1,1,t-1}\varepsilon_{1,t-1} + q_{1,2,t-1}\varepsilon_{2,t-1}\right)^2 \quad (18)$$

$$q_{1,2,t}q_{2,2,t} = q_{1,2}q_{2,2} + a_1q_{1,2,t-1}a_2q_{2,2,t-1} + 2b_1b_2q_{2,2,t-1}\varepsilon_{2,t-1}\left(q_{1,1,t-1}\varepsilon_{1,t-1} + q_{1,2,t-1}\varepsilon_{2,t-1}\right) \quad (19)$$

$$q_{2,2,t}^2 = q_{2,2}^2 + a_2^2q_{2,2,t-1}^2 + b_2^2\left(q_{2,2,t-1}\varepsilon_{2,t-1}\right)^2 \quad (20)$$

where $q_{1,1}, q_{1,2}, q_{2,2}, a_1, a_2, b_1, b_2$ are the parameters for the variance covariance system. A useful property of the linear MV-GARCH models is that the long run covariance matrix Σ has an analytic representation, see A for derivation. For this simple case the elements of the long run covariance matrix are computed by the following three equations

$$\tilde{q}_{1,1}^2 + \tilde{q}_{1,2}^2 = \left(q_{1,1}^2 + q_{1,2}^2\right)\left(a_1^2 + b_1^2\right)^{-1} \quad (21)$$

$$\tilde{q}_{1,2}\tilde{q}_{2,2} = q_{1,2}q_{2,2}\left(a_1a_2 + b_1b_2\right)^{-1} \quad (22)$$

$$\tilde{q}_{2,2}^2 = q_{2,2}^2\left(a_2^2 + b_2^2\right)^{-1} \quad (23)$$

where $\tilde{q}_{1,1}^2 + \tilde{q}_{1,2}^2, \tilde{q}_{1,2}\tilde{q}_{2,2}, \tilde{q}_{2,2}^2$ are the elements of the long run covariance matrix, with an upper triangular Cholesky factor containing the elements $\tilde{q}_{1,1}, \tilde{q}_{1,2}$ and $\tilde{q}_{2,2}$, which maybe exactly identified by sequential substitution from the third equation into the second and finally the first. For this model the solution is as follows,

$$\tilde{q}_{1,1} = \left(\left(q_{1,1}^2 + q_{1,2}^2\right) \left(a_1^2 + b_1^2\right)^{-1} - \frac{q_{1,2}q_{2,2}\left(a_1a_2 + b_1b_2\right)^{-1}}{(q_{2,2}^2(a_2^2 + b_2^2)^{-1})^{0.5}} \right)^{0.5} \tag{24}$$

$$\tilde{q}_{1,2} = \frac{q_{1,2}q_{2,2}\left(a_1a_2 + b_1b_2\right)^{-1}}{(q_{2,2}^2(a_2^2 + b_2^2)^{-1})^{0.5}} \tag{25}$$

$$\tilde{q}_{2,2} = (q_{2,2}^2(a_2^2 + b_2^2)^{-1})^{0.5} \tag{26}$$

Next, we set $\varepsilon_{i\in\{1,2\},t-k} = 0, \forall k \in \mathbb{N}$, which in this case reduces to $\varepsilon_{1,t-1} = \varepsilon_{2,t-1} = 0$. We can solve for $q_{1,1,t-1}, q_{1,2,t-1}$ and $q_{2,2,t-1}$, although in this specification it is redundant as the autoregression is of order 1. For completeness we can see this is solved by sequential substitution of the following:

$$\left(q_{1,1}^2 + q_{1,2}^2\right) \left(a_1^2 + b_1^2\right)^{-1} = q_{1,1}^2 + q_{1,2}^2 + a_1^2 \left(q_{1,1,t-1}^2 + q_{1,2,t-1}^2\right) \tag{27}$$

$$q_{1,2}q_{2,2}\left(a_1a_2 + b_1b_2\right)^{-1} = q_{1,2}q_{2,2} + a_1q_{1,2,t-1}a_2q_{2,2,t-1} \tag{28}$$

$$q_{2,2}^2\left(a_2^2 + b_2^2\right)^{-1} = q_{2,2}^2 + a_2^2q_{2,2,t-1}^2 \tag{29}$$

starting with $\mathrm{Re}((q_{2,2}^2a_2^{-2} + q_{2,2}^2b_2^{-2} - q_{2,2}^2)a_2^{-2})^{0.5} = q_{2,2,t-1}$. The usefulness of this specification is that there is a unique solution to $L\Sigma_t$ when $\Sigma_t = \Sigma$ and this does not require the imposition of any functional constraints, this is not the case with MV-GARCH specifications of order greater than one and for more complex first order models.

For brevity we shall restrict our example to the case when $\delta = [0, 1]'$, i.e. a single unit shock to $\varepsilon_{2,t}$. The long run expectation of $y_{1,t}$ and $y_{2,t}$ in this case is zero, therefore setting $y_{1,t-1} = 0$ and $y_{2,t-1} = 0$ the adjustment at time t will be

$$y_{1,t}^{\delta} = \frac{q_{1,2}q_{2,2}\left(a_1a_2 + b_1b_2\right)^{-1}}{(q_{2,2}^2(a_2^2 + b_2^2)^{-1})^{0.5}} \tag{30}$$

$$y_{2,t}^{\delta} = (q_{2,2}^2(a_2^2 + b_2^2)^{-1})^{0.5} \tag{31}$$

again it is important to emphasize that due to the ordering effect intrinsic in a Cholesky system both equations are shocked by a single unit innovation to $\varepsilon_{2,t}$ but only the first equation is shocked by an impulse in $\varepsilon_{1,t}$. We shall discuss an example with a non-upper triangular assumption on the structure of Q_t in the empirical example section.

We can recover the magnitude of the initial innovation by multiplication of $[y^{\delta}_{1,t}, y^{\delta}_{2,t}]'$, by multiplication by $(Q^{\delta}_t)^{-1} = Q^{-1}_t$. Of course this is unity for $\tilde{\delta}_{2,t}$, however $\tilde{\delta}_{1,t}$ is now

$$\begin{aligned}\tilde{\delta}_{j=1,i=2,s=0} &= \frac{q_{1,2}q_{2,2}\left(a_1a_2+b_1b_2\right)^{-1}}{(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5}} \\ &\quad \left(\left(q^2_{1,1}+q^2_{1,2}\right)\left(a^2_1+b^2_1\right)^{-1} - \frac{q_{1,2}q_{2,2}(a_1a_2+b_1b_2)^{-1}}{(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5}}\right)^{-0.5} \\ &\quad - \frac{q_{1,2}q_{2,2}(a_1a_2+b_1b_2)^{-1}}{(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5}\left(\left(q^2_{1,1}+q^2_{1,2}\right)\left(a^2_1+b^2_1\right)^{-1} - \frac{q_{1,2}q_{2,2}(a_1a_2+b_1b_2)^{-1}}{(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5}}\right)^{0.5}}.\end{aligned} \tag{32}$$

where $\tilde{\delta}_{j,i,s}$ is indexed by the j element in response to a shock in the i element at time step s. We can think of this as being the size of shock to $\varepsilon_{1,t}$ that would be equivalent to a single unit shock in $\varepsilon_{2,t}$. An important observation is that this need not be less than unity, therefore a shock in $\varepsilon_{2,t}$ can result in a larger shock in $\tilde{\delta}_{1,i=2}$. The innovation at step one is easily verified to be

$$y^{\delta}_{1,t+1} = \pi_{1,1}\frac{q_{1,2}q_{2,2}\left(a_1a_2+b_1b_2\right)^{-1}}{(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5}} + \pi_{1,2}(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5} \tag{33}$$

$$y^{\delta}_{2,t+1} = \pi_{2,1}\frac{q_{1,2}q_{2,2}\left(a_1a_2+b_1b_2\right)^{-1}}{(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5}} + \pi_{2,2}(q^2_{2,2}(a^2_2+b^2_2)^{-1})^{0.5} \tag{34}$$

recalling that in this specification $y^0_{1,t+1} = y^0_{2,t+1} = 0$. Once again we can multiply $[y^{\delta}_{1,t+1}, y^{\delta}_{2,t+1}]'$ by Q^{-1}_t, to compute $\tilde{\delta}_{j,i,s=1}$. This is simplified to

$$\tilde{\delta}_{j=1,i=2,s=1} = \tilde{q}^{-1}_{1,1}y^{\delta}_{1,t+1} - \frac{\tilde{q}_{1,2}y^{\delta}_{2,t+1}}{\tilde{q}_{1,1}\tilde{q}_{2,2}} \tag{35}$$

$$\tilde{\delta}_{j=2,i=2,s=1} = \tilde{q}^{-1}_{2,2}y^{\delta}_{2,t+1} \tag{36}$$

either of the previous two specifications will be asymptotically identical to Cholesky or generalised impulse response computed in the standard manner (although they are bias corrected for finite samples by capturing the MV-GARCH process in the disturbance term). However, these responses do not incorporate the expected change in the variance-covariance matrix. Incorporation of this component will give the econometrician (or policy maker) an indication of how variance-covariance shocks will enter the system of equations and comparison with the standard impulse response function. In certain cases this function will have a specific purpose for instance if the VAR is describing a vector of assets this plot will illustrate the expected change in the Sharp ratio given a shock to a single asset.

This adjusted impulse response is specified by the following system of equations updating the variance covariance matrix,

$$\left(\tilde{q}^{\delta_{i=2}}_{1,1,s=1}\right)^2 + \left(\tilde{q}^{\delta_{i=2}}_{1,2,s=1}\right)^2 = q^2_{1,1} + q^2_{1,2} + a^2_1 \left(\tilde{q}^2_{1,1} + \tilde{q}^2_{1,2}\right) + b^2_1 \left(\tilde{q}_{1,1}\tilde{\delta}_{1,1,0} + \tilde{q}_{1,2}\tilde{\delta}_{2,1,0}\right)^2 \quad (37)$$

$$\tilde{q}^{\delta_{i=2}}_{1,2,s=1}\tilde{q}^{\delta_{i=2}}_{2,2,s=1} = q_{1,2}q_{2,2} + a_1 q_{1,2,t-1} a_2 q_{2,2,t-1} + 2b_1 b_2 \tilde{q}_{2,2}\tilde{\delta}_{2,1,0}\left(\tilde{q}_{1,1}\tilde{\delta}_{1,1,0} + \tilde{q}_{1,2}\tilde{\delta}_{2,1,0}\right) \quad (38)$$

$$\left(\tilde{q}^{\delta_{i=2}}_{2,2,s=1}\right)^2 = q^2_{2,2} + a^2_2 q^2_{2,2,t-1} + b^2_2\left(\tilde{q}_{2,2}\tilde{\delta}_{2,1,0}\right)^2 \quad (39)$$

where we index the innovations $\tilde{q}^{\delta_i}_{j,k,s}$, by j, k element response to as shock in the i element of ε_t at the s step. Therefore the adjusted response at step one is

$$\tilde{y}^{\delta_{i=2}}_{j=1,s=1} = \tilde{\delta}_{1,2,1}\tilde{q}^{\delta_2}_{1,1,1} + \tilde{\delta}_{2,2,1}\tilde{q}^{\delta_2}_{1,2,1} \quad (40)$$

$$\tilde{y}^{\delta_{i=2}}_{j=2,s=1} = \tilde{\delta}_{2,2,1}\tilde{q}^{\delta_2}_{2,2,1} \quad (41)$$

We can then substitute $\tilde{y}^{\delta_{i=2}}_{j=1,s=1}$ and $\tilde{y}^{\delta_{i=2}}_{j=2,s=1}$ for $y_{1,t}$ and $y_{2,t}$ to compute the next update at $s = 2$. Similarly we can update $\tilde{q}^{\delta_{i=2}}_{1,1,s=2}$, $\tilde{q}^{\delta_{i=2}}_{1,2,s=2}$ and $\tilde{q}^{\delta_{i=2}}_{2,2,s=2}$ using $\tilde{\delta}_{1,2,1}$ and $\tilde{\delta}_{2,2,1}$. A useful aspect of the VAR-BEKK type approach is that at each step the left hand side of the equations is computed purely from expressions of the parameters $\pi_{1,1}, \pi_{1,2}, \pi_{2,1}, \pi_{2,2}, q_{1,1}, q_{1,2}, q_{2,2}, a_1, a_2, b_1$ and b_2. If the parameters are estimated with error then we can quickly construct error bounds by drawing new sets of parameters using the parameter variance covariance matrix. Increasing the lag length and dimensionality of the system results in a quadratic increase in computations.

For standard VAR models, when $w = 0$, the approach above is useful in augmenting existing impulse responses to illustrate the impact of shocks on the disturbance process. However, in a structural VAR (SVAR) context, the decomposition of $\Sigma_t = Q_t Q_t'$ has a meaning in terms of the choice of decomposition. SVAR models impose from theory the possibility that the elements of the vector y_t have some instantaneous effect on each other. In the most general sense, this is written out as a projection of the following form $\tilde{\Gamma} y_{t+1} =$

$\tilde{\Phi} L \tilde{v}_t + \tilde{\epsilon}_t$, where $\tilde{\Gamma}$ is a matrix or vector function such that $\tilde{\gamma} : \mathbb{R}^n \to \mathbb{R}^n$. In this case the covariance matrix will be decomposed as follows,

$$\Sigma_t = Q_t Q_t' = \tilde{\Gamma}^{-1} \tilde{\Omega}_t \tilde{\Gamma}'^{-1} \tag{42}$$

where $\tilde{\Omega}_t$ is time varying positive definite matrix. It is relatively straight forward to establish that by recomputing the structural moving average (SMA) process by filtering the MV-GARCH and then imposing the long run covariance matrix Σ and identifying $\tilde{\Gamma}$, then computing $\tilde{\Omega}$ and finally the impulse responses, is the same approach to the re-identification of $\tilde{\delta}_{i,j,s}$ and $\tilde{y}_{i,s}^{\delta_i}$ in the example with no structural restrictions other than the upper triangular for of Q_t.

3. Empirical Examples I: Estimated Models

The VAR-MV-GARCH model offers a great deal of flexibility by explicitly incorporating time varying second moments into the disturbance structure, these changes can be thought of as changes in risk premia or volatility adjustments following shocks. This section outlines two empirical applications for this type of model, the first is a simple analysis of the impulse responses between inflation and output, of inherent relevance to monetary policy. The second, outlines a structural VAR in the Blanchard-Quah mould and outlines a method identifying temporary and permanent shocks in the presence of MV-GARCH disturbances.

3.1. Inflation and Output

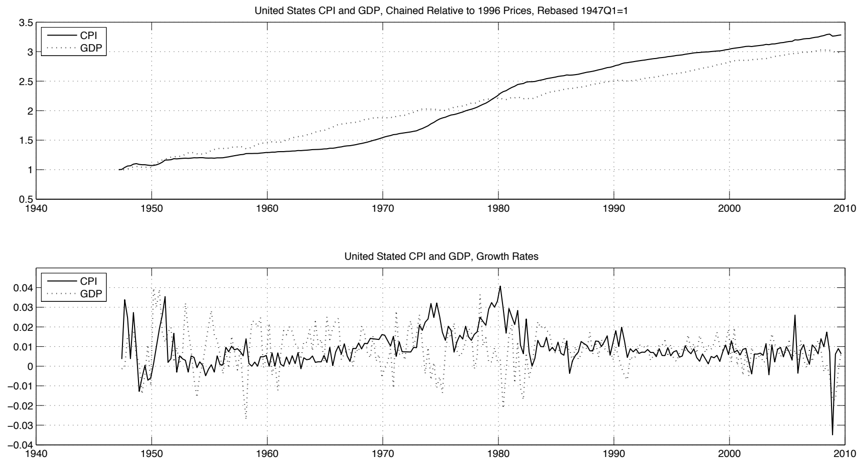
Our first example uses a simple unrestricted VAR-MV-GARCH to fit a model of inflation and output for quarterly post war data from the United States, in the vein of King and Watson (1994) and Bullard and Keating (1995). For this purpose we again use a BEKK model for the covariance process, see A, however we do not impose a restriction on the matrices A and B to be diagonal.

We estimate two forms of the model, first with MV-GARCH effects suppressed and second with MV-GARCH included. We report the likelihood ratio statistic

$$LR^* = 2\left(\mathfrak{L}^*\left(\hat{\theta}\right) - \mathfrak{L}^*\left(\hat{\theta}^r\right)\right), \tag{43}$$

where θ^r is the restricted model with the MV-GARCH model parameters suppressed, a restriction of $\nu = 0.5n(n+1) + pn^2 + qn^2$ free parameters. Therefore if $LR > c_\alpha$, then we reject the null of model equality, where c_α is $\chi^2(\nu)$, for a choice of significance α.

Figure 1: Plots of the Evolution of CPI and GDP for the US from 1947Q1 to 2010Q3. The upper plot shows the variables rebased such that 1947Q1=1. The GDP measure is computed using the chained GDP96 time series reported by the Federal Reserve Databank, St Louis. The lower plot shows the relative growth rates in each variable.



Setting $y_t = [\Delta U\,(CPI)\,, \Delta U\,(GDP)]'$, where $U(\cdot)$ is an $MA(p)$ univariate pre whitening filter. For the VAR lag order r we utilise the Bayesian Information Criterion (BIC). For the MV-GARCH lag order we use a mixture of BIC and likelihood ratio tests to discriminate the order of the MV-GARCH model.

The data is obtained from the Federal reserve databank, St Louis and runs for 251 quarters 1947Q1 to 2009Q3 and the CPI, GDP, inflation and growth rates are plotted in Figure 1. We have opted for a VAR(2)-MV-GARCH(1,1) specification using evidence from the BIC criterion[1]. Table 1 reports the regression results for the VAR-MV-GARCH model.

We test for the existence of MV-GARCH effects. The estimated log likelihood for the restricted model with MV-GARCH effects surpressed is 1429.87 and for the unrestricted model with MV-GARCH(1,1) disturbances is 1714.53. Twice the difference in the log likelihoods is 569.34, which is substantially higher than the critical statistic for $\alpha = 95\%$ significance, which is 19.6751, suggesting that the MV-GARCH model should be included instead of a static covariance model[2].

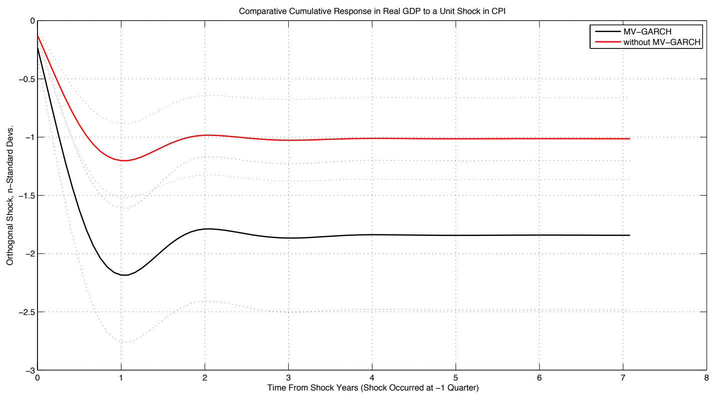
[1]Likelihood ratio tests suggest that a longer lag length in mean and variance could be up to VAR(4)-MV-GARCH(3,2)

[2]Full results for this model are available from the authors on request.

Table 1: The estimated model parameters for the VAR(2)-MV-GARCH(1,1) model specification for $y_t = [\Delta U(CPI), \Delta U(GDP)]'$. z-statistics are presented in brackets for each coefficient. i represents each equation for the VAR model. The disturbance model is presented as follows, $q_{1,1}$, $q_{1,2}$ and $q_{2,2}$ are the intercept coefficients from the matrix Q. $a_{1,1}, a_{1,2}, a_{2,1}$ and $a_{2,2}$ are the elements of the square matrix, A, of ARCH parameters $b_{1,1}, b_{1,2}, b_{2,1}$ and $b_{2,2}$ are the elements of the square matrix, B, of GARCH parameters from the BEKK model. Further ancillary diagnostics are available from the authors, however these results are well documented in the literature, see Bullard and Keating (1995) for a cross country comparison and our results, when stripped of the MV-GARCH specification, are in keeping with them.

	VAR(2)		MV-GARCH(1,1)	
	$\Delta \log CPI$	$\Delta \log GDP$		Estimates
μ_i	0.46368	0.10933	$q_{1,1}$	0.00134
	[8.17811]	[1.29177]		[2.50319]
$\pi_{1,i}$	0.08123	0.39674	$q_{1,2}$	0.00013
	[1.93222]	[6.32154]		[0.16217]
$\pi_{2,i}$	0.14441	-0.16634	$q_{2,2}$	0.0015
	[2.56976]	[-1.98275]		[0.89134]
$\pi_{3,i}$	0.06286	0.0817	$a_{1,1}$	0.58793
	[1.48139]	[1.28973]		[4.21564]
$\pi_{4,i}$	0.0011	0.00259	$a_{1,2}$	0
	[2.78373]	[4.36064]		[-0.00000]
			$a_{2,1}$	0
				[0.00000]
			$a_{2,2}$	0.52112
				[5.68957]
			$b_{1,1}$	0.77552
				[6.36025]
			$b_{1,2}$	0
				[-0.00000]
			$b_{2,1}$	0
				[-0.00000]
			$b_{2,2}$	0.64324
				[5.47102]

Figure 2: Comparison response in real GDP to a unit shock in ϵ_{CPI}. The impulse responses are presented in terms of numbers of standard deviations of the 'without MV-GARCH' model for appropriate comparison.



Comparison of IRFs with and without MV-GARCH

From the coefficients in Table 1 we see that $a_{2,1}$, $a_{1,2}$, $b_{2,1}$ and $b_{1,2}$ are zero, therefore imposition of the diagonal restrictions would not necessarily affect our calculations. The coefficients indicate that whilst there is a degree of cross persistence in variance, there is little persistence in covariance.

Following Bullard and Keating (1995) we report the cumulative impulse response functions for the estimated system. To compare the impact of the inclusion of MV-GARCH effects to a similar we conduct the following comparison. We have first estimated the system without MV-GARCH effects and compute the impulse response in mean using the covariance matrix Σ^* derived from the OLS estimator. We then compute the cumulative IRFs from the model with MV-GARCH effects included and plot them for comparison. Figure 2 plots the IRF of GDP growth to an inflation shock in number of standard deviations (chained to the original OLS model) with and without MV-GARCH effects. The mid point response shows that a cumulative shock to the VAR-MV-GARCH model is equivalent to double the number of standard deviations of the VAR IRFs when MV-GARCH effects are ignored.

Price shocks originating from the supply side are cumulatively far more damaging to output than would be previously estimated, when the variance effects are not taken into account. Both the amplitude and the duration of the shock are more pronounced as the influence of the original shock decays at a far slower rate because its variance has changed in subsequent periods. In this case the macroeconomic policy response required to restore the economy to equilibrium will require a substantial injection of demand to restore output to an acceptable level over an above that which would be expected from the without MV-GARCH model.

3.2. A Blanchard-Quah Approach with MV-GARCH Disturbances

In the methodology section we outlined the importance of the covariance decomposition for an SVAR type analysis. In a famous example Blanchard and Quah (1989) demonstrate a simple bivariate SVAR model of output and unemployment. The structural decomposition is designed to impute impulse response functions for shocks to supply and demand to the US economy. We will demonstrate the Blanchard and Quah (1989) model to include time varying second moments in the disturbance process.

Our analysis will be in two parts, first we establish whether MV-GARCH disturbances should be included in the model structure and then we demonstrate the adjustments to the impulse response functions when the MV-GARCH model is included, using our approach.

In our notation $y_t = [\Delta \log(GDP), \log(Unemployment)]'$. Via the companion matrix in Equation 3, for an centered estimated stationary VAR system we can interchange between the VMA and VAR representations, $y_t = \sum_{i=1}^{\infty} \hat{F}_i \hat{v}_{t-i} + \hat{v}_t$ and $y_t = \sum_{i=1}^{r} \hat{\Pi}_i y_{t-i} + \hat{u}_t$. Where the $\hat{\Pi}$ matrix is estimated via maximum likelihood estimation with potentially conditionally heteroskedastic disturbances and F_i is computed via the companion matrix in Equation 3. The estimated VMA and VAR shocks $\hat{v}_t$ and $\hat{u}$ respectively, are 'statistical' disturbances from the VAR model which are presumed to be linear combinations of the underlying structural disturbances. Let $\mathbb{E}(v_t v_t') = \Omega$, this is estimated by $T^{-1} \sum_{i=1}^{T} \hat{v}_t \hat{v}_t$.

The approach of Blanchard and Quah (1989), is relatively simple, specify another VMA process $y_t = \sum_{i=1}^{\infty} C_i e_{t-i} + C_0 e_t$, where the sequence of matrices $C_i = [c_{jk}]_{j,k \in \{1,2\}}$, is such that $\sum_{i=1}^{\infty} c_{11,i} = 0$. Implying that the shock $e_{1,t}$ has no long run impact on the level of GDP itself. The sequence C_i is given by the identification $A_i = F_i C_0$. To identify C_0 is relatively simple as by construction $A_0 A_0' = \Omega$, from the previous restriction on the upper left $\sum_{i=1}^{\infty} a_{11,i} = 0$ the matrix is explicitly identified see Blanchard and Quah (1989) for the full proof. If we assume that the variance process of u_t is BEKK type autoregressive process

$$\Sigma_t = QQ' + \sum_{j=1}^{q} A_j u_{t-1} u'_{t-1} A_j + \sum_{i=1}^{p} B_1 \Sigma_{t-1} B'_1 \tag{44}$$

Table 2: Decomposition of Shocks

Shock	Designation	Derived From	Covariance
Observed	u_t	$y_t - \mathbb{E}(y_t)$	Σ_t
Structural cholesky	ε_t	$u_t \Sigma_t^{-\frac{1}{2}}$	I
Vector Moving Average	v_t	$\Sigma^{\frac{1}{2}} \varepsilon_t$ and Π via matrix Φ	$\Omega = C_0 C_0{}'$
Structural shocks	$e_t = y_t - \sum_{i=1}^{\infty} C_i e_{t-i}$	filtered from $C_i = F_i C_0$	Ω

We can then filter the conditionally heteroskedastic disturbances and recompute the structural moving average process A_i at each point, with an updated Ω. This representation has several advantages in allowing the model to map to more complex shock responses, whilst maintaining the long run economically significant identification. For the BEKK model the long run restriction is simple to impose because the long run covariance is analytically defined as

$$vec\,(\Sigma) = \left(I - \sum_{i=1}^{p} (A_i \otimes A_i)' - \sum_{i=1}^{p} (B_i \otimes B_i)' \right)^{-1} vec\,(QQ') \tag{45}$$

The impulse responses are then recomputed by plotting the elements of the updated $\hat{A}_i$. Table 2 describes the identification of the various shocks within the model. The model is estimated over quarterly data for the US from the Federal Research Economic Data (FRED) database from 1948Q1 to 2010Q1, GDP chained at 2005 prices (FRED code GDPC96) and unemployment is the total number of unemployed persons of over 16 years of age in the US, seasonally adjusted (FRED code UNEMPLOY).

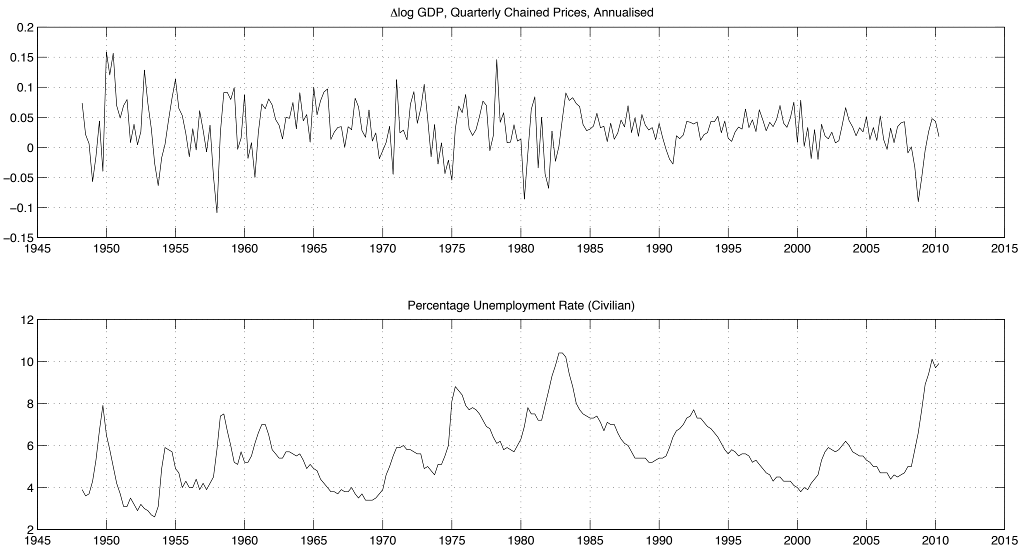
The responses of output and unemployment to supply (permanent) and demand (temporary) shocks are presented in Figures 4 and 5. The IRFs are calculated using two versions of the structural two equation VAR model. In the first specification we select the model without MV-GARCH effects using the optimal number of lags indiciated by the BIC. This model incorporates eight lags. The second model, which allows for MV-GARCH effects has four lags under the same selection criterion. Table 3 presents the estimated coefficients for both models and the resulting log likelihoods. The models are non-nested and therefore cannot be directly compared using a likelihoopd ratio. However, it is noted that the VAR with MV-GARCH disturbances, whilst having five parameters less than the simple linear model, has a higher log likelihood.

The response of output to shocks in demand when MV-GARCH effects are taken into

Table 3: The table below presents the comparative results for the best VAR with and without MV-GARCH disturbances. Using the Bayesian Information Criterion (BIC) the optimal lag length for the VAR without MV-GARCH disturbances is eight lags. With an MV-GARCH(1,1) disturbance process this is reduced to four lags. BIC analysis suggests that the MV-GARCH order should be 1,1 or 2,1, we have decided to use the most parsimonious model, from likelihood ratio testing, which is the MV-GARCH(1,1). The log likelihood of the VAR(8) without MV-GARCH disturbances is 639.1135, whilst for the VAR(4)-MV-GARCH(1,1) it is 652.2543. The VAR(8) model has 34 parameters, whilst the VAR(4)-MV-GARCH(1,1) actually reduces the total parameters to 29, 18 for the VAR model and 11 for the MV-GARCH model. The models are now non-nested and so likelihood ratio testing is not possible. Addition of the second GARCH term would require 4 extra parameters, however testing suggested that IRF profiles are indistinguishable, t-statistics for the VAR(8) model and z-statistics for the VAR(4)-MV-GARCH(1,1) model are presented in square brackets.

	VAR(8)		VAR(4)-MV-GARCH(1,1)			
	$\Delta \log GDP$	$\log UNEMP$	$\Delta \log GDP$	$\log UNEMP$	MV-GARCH	
μ_i	0.34996	-0.05195	0.30371	-0.04376	$q_{1,1}$	0.16252
	[4.77374]	[-8.45874]	[4.07737]	[-6.60084]		[0.72299]
$\pi_{1,i}$	0.33968	0.95782	0.34446	0.97472	$q_{1,2}$	0.04233
	[0.39765]	[13.38389]	[0.40889]	[12.99962]		[1.39964]
$\pi_{2,i}$	0.16444	-0.03737	0.23318	-0.04643	$q_{2,2}$	-0.00009
	[2.00700]	[-5.44508]	[2.90411]	[-6.49709]		[-0.00311]
$\pi_{3,i}$	0.73499	-0.05159	1.35592	-0.14544	$a_{1,1}$	-0.13082
	[0.63577]	[-0.53266]	[1.14428]	[-1.37903]		[-1.66345]
$\pi_{4,i}$	-0.0252	-0.01551	0.0266	-0.01914	$a_{1,2}$	0.04272
	[-0.29215]	[-2.14729]	[0.30619]	[-2.47538]		[3.63805]
$\pi_{5,i}$	-0.29918	0.00848	-0.52098	0.02238	$a_{2,1}$	-2.83001
	[-0.25808]	[0.08732]	[-0.44408]	[0.21435]		[-5.00545]
$\pi_{6,i}$	0.15165	-0.00796	0.03177	-0.00293	$a_{2,2}$	0.43955
	[1.78585]	[-1.11989]	[0.41440]	[-0.43042]		[4.72539]
$\pi_{7,i}$	2.20815	-0.07769	-0.83504	0.10136	$b_{1,1}$	0.97659
	[1.91186]	[-0.80295]	[-1.07705]	[1.46886]		[79.77988]
$\pi_{8,i}$	-0.00094	0.0149	-0.68955	0.65298	$b_{1,2}$	-0.00001
	[-0.01097]	[2.07429]	[-0.58148]	[6.18647]		[-0.00810]
$\pi_{9,i}$	-3.16814	0.32383			$b_{2,1}$	0.04578
	[-2.73135]	[3.33251]				[0.22633]
$\pi_{10,i}$	0.14621	-0.00633			$b_{2,2}$	0.87702
	[1.70848]	[-0.88388]				[14.27936]
$\pi_{11,i}$	0.52568	-0.24163				
	[0.44498]	[-2.44144]				
$\pi_{12,i}$	0.04319	-0.00264				
	[0.50429]	[-0.36882]				
$\pi_{13,i}$	1.08702	0.04753				
	[0.94084]	[0.49111]				
$\pi_{14,i}$	0.03418	0.00022				
	[0.46633]	[0.03727]				
$\pi_{15,i}$	-1.08325	-0.01835				
	[-1.41081]	[-0.28540]				
$\pi_{16,i}$	-1.62397	0.65566				
	[-1.04431]	[5.03272]				

Figure 3: Time series plots of the data utilised in the Blanchard-Quah with MV-GARCH effects. The upper plot is for GDP chained



account are more pronounced compared to the same IRF without MV-GARCH effects. Both the scale and evolution of output when GARCH effects are taken into account indicate its response to demand shocks is both substantial and longer lasting than previously thought when such conditional variance effects are not taken into account. The confidence intervals overlap rendering the exact statistical seperation of these responses problematic. Nevertheless the point estimate of the 'with MV-GARCH' IRF always lies always above the point estimate of the 'no-MV-GARCH' response indicating that the expectation that demand shocks have substantial influence in output when 'MV-GARCH' effects are included.

The permanent supply shocks, with and without MV-GARCH effects are presented in the lower part of Figure 4. Unlike the previous case of demand shocks, here we are able to establish a clear statistical distinction between the responses of the two models as the confidence intervals of the resulting IRFs are distinct. The inclusions of MV-GARCH results in substantially higher output response compared to the usual VAR model.

The corresponding IRFs of unemployment to supply shocks for both models are presented in the lower part of Figure 5. The responses are very similar indeed both in terms of scale and time evolution. As the corresponding confidence interval exhibit a very substantial overlap it is impossible to provide a clear statistical separation between the responses of the

Figure 4: Comparative IRF plots for the responses in output to supply and demand shocks. The plot for the 'with MV-GARCH' models are scaled so that they are in terms of 'without MV-GARCH' standard deviations for direct comparison. The error bounds are computed using 10,000 draws from the parameter distribution.

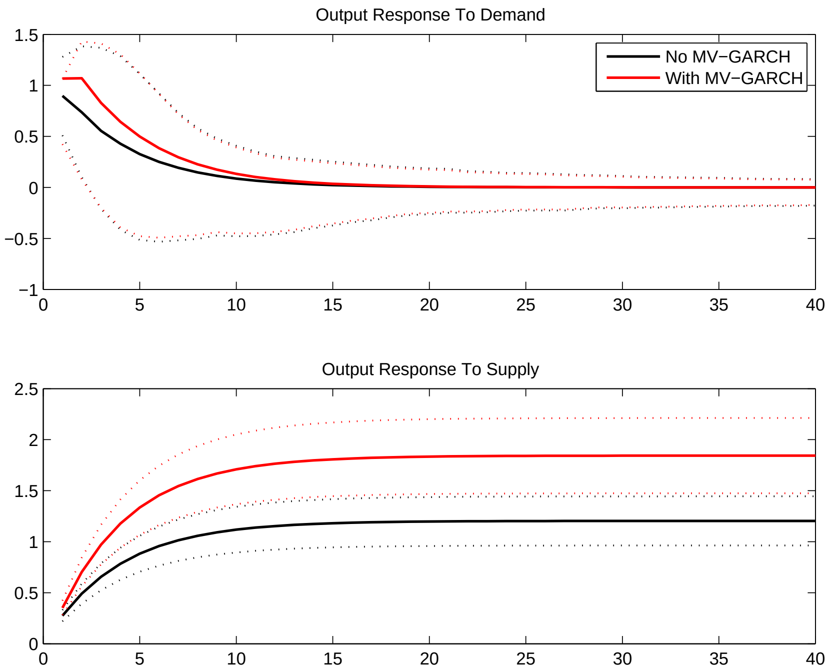
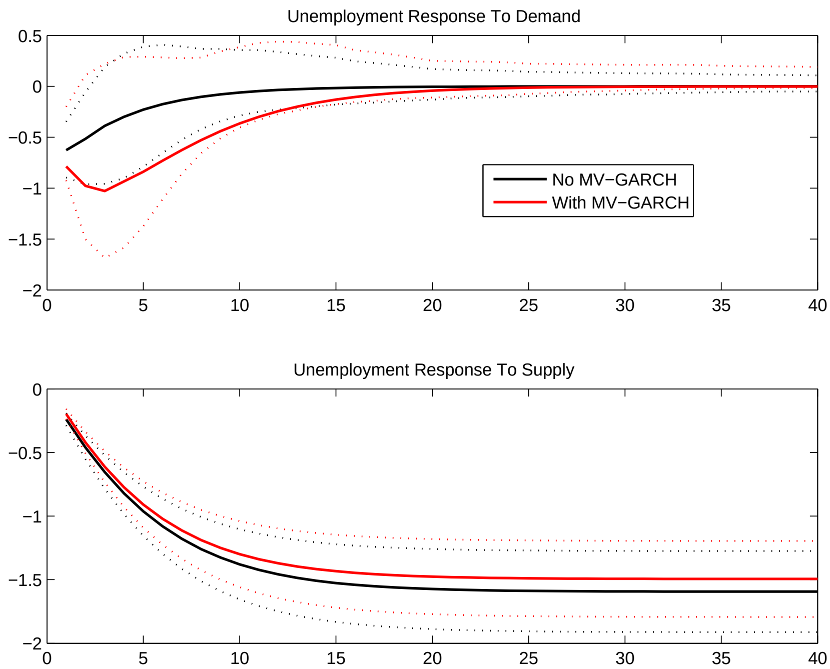


Figure 5: Comparative IRF plots for the responses in unemployment to supply and demand shocks. The plots are produced using the same procedure for the output plots.



two models. By and large our results show that the reduction in unemployment is somewhat less pronounced when the conditional variance effects are included. This result is in direct contrast to our findings regarding output, where the differential response of output is well defined and more importantly the impact on output from a supply shock is more pronounced in the 'with MV-GARCH' model.

The impact of the demand shock, in both cases results in unemploymnet reductions although the time evolution of unemployment following the shock differs substantially between the models. Under the conventional model, unemployment declines in the first period and subsequentialy rises monotonicaly to its original level. When MV-GARCH effects are included the initial reduction in unemployment is a greater magnitude (-0.6 compared to -0.75) and more importantly this decline continues over the next (two/three) periods achieving the maximum reduction of -1 standard deviations of the 'without MV-GARCH' umemployment equation. Although there is a substantial overlap between the corresponding confidence intervals , there is sufficient distinction in the early periods to suggest that overall the influence of the aggregare demand shock on unemployment is more pronounced in the presence of MV-GARCH effects. An important point to note is that the lower 95% confidence interval of the impulse response function for the 'with MV-GARCH' model drops to nearly -1.5 standard deviations. The implication of this is that the risks to unemployment following a negative demand shock are far more pronounced when the conditional variance effects are included.

Such findings can account partially for the developmemts of output and unemployment in the USA during the period after the 2007-2010 financial crisis. The sharp fall in demand resulting from the dramatic reduction in the availability of credit has reduced output and increased unemployment but whilst output has recovered quite rapidly unemployment has continued to rise. Following a negative demand shock policy responses that ignore the variance dynamics will underestimate the both the increase in unemployment and its future evolution.

4. Empirical Example II: A Calibrated Model

We now apply our methodology to VAR(1) models that have been derived from the analytical micro-foundations of Dynamic Stochastic General Equilibrium (DSGE) models. These models are now very well established in the evaluation of economic policy. In this section we shall apply calibration parameters from the extant literature and we will the introduce alternative models of vector conditional heteroskedasticity and compare the resulting structural impulse response functions under alternative policy shocks.

Such models have VAR representations only under very specific conditions, and their identification depends upon the manner that the vector of "deep parameters" (θ) which characterize the structural model are trasnsformed into the coefficients of the VAR, Komunjer and Ng (2011), whose notation we will be using, derive the conditions for the dynamic identification of the model parameters from the first two moments of the data available for the 'observed' variables. Denoting by X_{t+1} the state variables and by Y_{t+1} the vector of the observed variables, and by ε_{t+1}, the vector of structural distrurbances; the system's state-space representation is given by:

$$X_{t+1} = A(\theta)X_t + B(\theta)\varepsilon_{t+1} \tag{46a}$$

$$Y_{t+1} = C(\theta)X_t + D(\theta)\varepsilon_{t+1} \tag{47}$$

Under the joint assumptions of stationarity and the existence of a positive definite long-run variance-covariance matrix of the structural distrurbances, this specification admits a number of functional forms for the conditional/short-run covariance matrix, such as GARCH and stochastic volatility. Justiniano et al. (2011) allow for the existence of stochastic volatility in a DSGE model and attempt to explain its time evolution over the period of the Great Moderation. Benigno et al. (2011) find that in the context of an open economy shocks in the volatility of nominal and real shocks can account for the observed persistent deviation of the nominal exchange rate from its value predicted by economic theory.

Our interest in this case is to examine the impact of the time varying volatility of the structural disturbances on the shape and magnitude of the impulse response functions of the observed variables following an innovation and compare it to the case where the error volatility is set permanently to its long run unconditional value. To this end we use the seven equation DSGE model developed by An and Schorfheide (2007), which is a standard test case in the literature. This model consists of four observed variables (the policy interest rate, output, inflation and observed consumption) and four structural disturbances arising from shock to a) technology, b) government spending, c) monetary policy and finally d) errors arising between the "true" and observed consumption.

The model requires the calibration of 9 parameters which link the seven endogenous variables and a number of their expectations contained in the model. The equations of the state variables and shocks are given by the following system:

$$
\begin{aligned}
y_t &= \mathbb{E}_t(y_{t+1} + g_t - \mathbb{E}_t(g_{t+1}) - (1/\tau)(r_t - \mathbb{E}_t\pi_{t+1} - \mathbb{E}_t(z_{t+1}) && (48)\\
\pi_t &= \beta\mathbb{E}_t(\pi_{t+1}) + \kappa(y_t - g_t) && (49)\\
c_t &= y_t - g_t && (50)\\
r_t &= \psi_\pi\pi_t + \psi_y(y_t - g_t) + \varepsilon_{rt} && (51)\\
g_t &= \rho_g g_{t-1} + \varepsilon_{gt} && (52)\\
z_t &= z_{t-1} + \varepsilon_{zt} && (53)\\
c_t^\varepsilon &= c_t + \varepsilon_{ct} && (54)
\end{aligned}
$$

where the variables are y_t is output, π is inflation, z_t is total factor productivity, g_t is government spending, r_t is the real interest rate and c_t is true consumption.

The mode itself is characterized by a set of ten of parameters such as: τ is the inverse intertemporal elasticity of substitution. β is the global discount factor. ν is the inverse elasticity of demand. ϕ is the degree of price stickiness. Π is the steady-state inflation rate. ψ_π is the Taylor-rule coefficient on the inflation gap (aggressiveness of monetary policy). ψ_y is the Taylor-rule coefficient on the output gap. ρ_z and ρ_g are the mean equation persistence in technology and government expenditure. Finally, $\sigma_z, \sigma_g, \sigma_r$ and σ_c are the standard deviations for the shocks to technology (ε_{zt}), government expenditure (ε_{gt}), monetary policy (ε_{rt}) and observed consumption (ε_{ct}) respectively.

Morris (2012b) and Morris (2012a) demonstrate for, respectively, three and four equation models that provided that when the number of observed variables is greater or equal to the number state variables the structural model can be rewritten in terms of a fully identified VAR(1) model as follows:

$$\begin{bmatrix} r_t \\ y_t \\ \pi_t \\ c_t^\epsilon \end{bmatrix} = \begin{bmatrix} \frac{\omega_1\psi_\pi\rho_z}{\omega 3+\omega 1\psi_\pi} & 0 & \frac{\omega_1\omega_3\rho_z}{\kappa(\omega_3+\omega_1\psi_\pi)} & 0 \\ \psi_\pi\frac{(1-\rho_g)\beta_z+\beta_z\rho_z}{\omega_3+\omega_1\psi_\pi} & \rho_g & \frac{\beta_z\omega_3(\rho_z-\rho_g)-\psi_\pi\omega_1\rho_g}{\kappa(\omega_3+\omega_1\psi_\pi)} & 0 \\ \frac{\kappa\psi_\pi\rho_z}{\omega_3+\omega_1\psi_\pi} & 0 & \frac{\omega_3\rho_z}{\omega_3+\omega_1\psi_\pi} & 0 \\ \frac{\beta_z\psi_\pi\rho_z}{\omega_3+\omega_1\psi_\pi} & 0 & \frac{\beta_z\omega_3\rho_z}{\kappa(\omega_3+\omega_1\psi_\pi)} & 0 \end{bmatrix} \begin{bmatrix} r_{t-1} \\ y_{t-1} \\ \pi_{t-1} \\ c_{t-1}^\epsilon \end{bmatrix}$$
$$+ \begin{bmatrix} \omega_1 & 0 & \omega_3 & 0 \\ \beta_z & 1 & -\psi_\pi & 0 \\ \kappa & 0 & -\kappa\psi_\pi & 0 \\ \beta_z & 0 & -\psi_\pi & 1 \end{bmatrix} \begin{bmatrix} \frac{\rho_z}{\omega_2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{1}{\tau+\psi_\kappa} & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{zt} \\ \varepsilon_{gt} \\ \varepsilon_{rt} \\ \varepsilon_{ct} \end{bmatrix} \tag{55}$$

Given the following ancillary parameters $\beta_z = 1 - \rho_z\beta$; $\beta_\Pi = 1 - \pi\beta$; $\psi_\kappa = \psi_y + \kappa\psi_\pi$; $\omega_1 = \beta_z\psi_y + \kappa\psi_\pi$; $\omega_2 = \beta_z((1-\rho_z)\tau + \psi_\kappa) - \kappa\beta_\pi$ and $\omega_3 = \tau + (1-\psi_\pi)\psi_\kappa$, where all the elements in both matrices are functions of the calibrated structural parameters. The system's variance-covariance matrix is given by $\Omega = (D\Sigma_\varepsilon D')$ where Σ_ε is a diagonal matrix whose elements are the unconditional long-run volatilities of the structural errors. In this case, the matrix D is given by:

$$D = \begin{bmatrix} \omega_1 & 0 & \omega_3 & 0 \\ \beta_z & 1 & -\psi_\pi & 0 \\ \kappa & 0 & -\kappa\psi_\pi & 0 \\ \beta_z & 0 & -\psi_\pi & 1 \end{bmatrix} \begin{bmatrix} \frac{\rho_z}{\omega_2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{1}{\tau+\psi_\kappa} & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{56}$$

For our purposes we use the point parameter calibrations suggested by Morris (2012a), which are described as follows:

$$\begin{array}{ccccccccccccc} \tau & \beta & \nu & \phi & \Pi & \psi_\pi & \psi_y & \rho_z & \rho_g & \sigma_z & \sigma_r & \sigma_g & \sigma_c \\ 2 & 0.9975 & 0.1 & 53.68 & 1.008 & 0.08 & 0.01 & 0.9 & 0.95 & 0.3 & 0.2 & 0.6 & 0.5 \end{array} \tag{57}$$

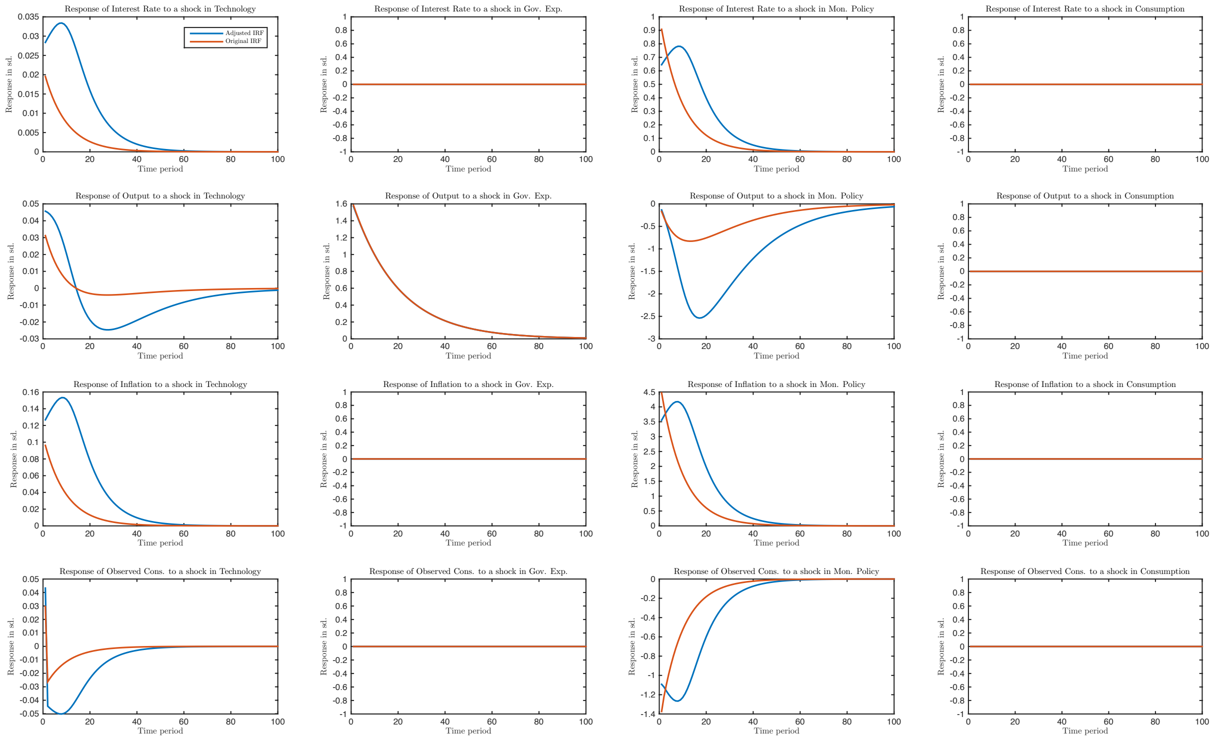
4.1. Introducing Diagonal ARCH/GARCH Effects

We will now calculate the impulse response functions of this model in the presence of monetary policy an productivity shocks under different conditions of dynamic covariance evolution. Subsequently, we will compare these impulse response functions to the impulse response functions when such GARCH effects are ignored. We introduce such effects to interest rates and technology. There is ample evidence that interest rates in the USA and Europe do exhibit significant ARCH effects , which have been well established in the empirical literature, see Engle et al. (1987) and Grey (1996). There is also accumulating evidence that the standard deviation of innovations to TFP are time dependent see Asker et al. (2012) and Pancrazi and Vukotic (2011) for a discussion of the methodologies in computing such measures. These are by no means the only GARCH specifications which can be accommodated in this model, but our purpose is to demonstrate within the confines of a relatively simple covariance structure, using two parameters to be calibrated , the impact of the presence of the dynamic covariance on the system's IRF following a shock. We constrain the long run covariance to be diagonal and attain the values postulated above. Assuming that the covariance process is near integrated, then the dynamic structure depends upon two parameters only as the stability conditions require that the squares of the diagonal ARCH (a_i) and GARCH (b_i) parameters sum up to near unity. Hence $b_i = \lambda/(\lambda+1)\sqrt{1-a_i^2}$, for an arbitrarily large number λ, for which we will choose 99. As the covariance process is of the BEKK type, we impose the following functional form on the A and B first order parameter matrices:

$$A = \begin{bmatrix} a_z & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a_r & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \frac{99}{100}\sqrt{1-a_z^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{99}{100}\sqrt{1-a_r^2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The BEKK model generally constrains us to consider only a relatively small number of cases if we impose stationarity on the covariance process. As such, we can generate two main types of effect. First, the GARCH autoregressive coefficient is relatively large compared to

Figure 6: Case 1: Low ARCH high persistence $a_z = a_r = 0.5$



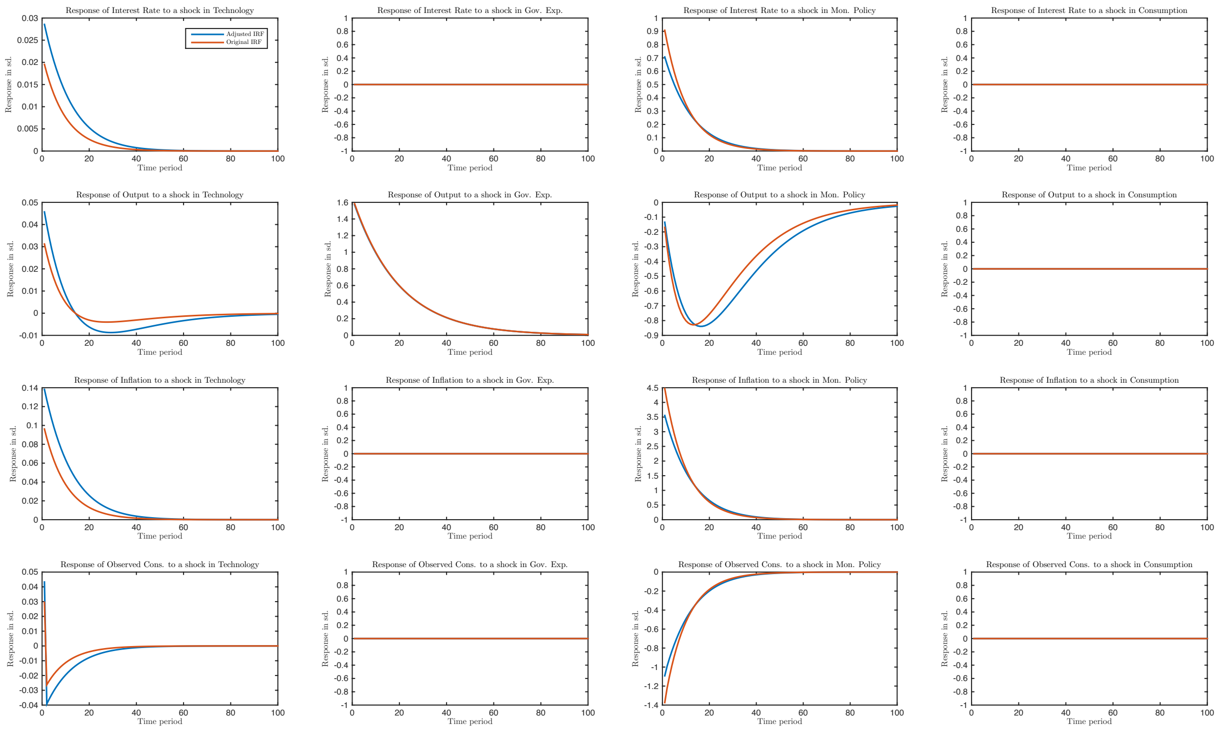
the ARCH coefficient $a_{i\in\{z,r\}} \leq 0.5$; in this case the GARCH effects are more persistent, however, the direct transmission of a shock from $\varepsilon_{i\in\{z,r\}}$ to $\sigma_{i\in\{z,r\}}$ is not as pronounced as in our next case. Second, when the ARCH effects dominate, the transmission of shocks is high, but the overall level of persistence is low, for example when $a_{i\in\{z,r\}} \geq 0.8$. As $a_i \to 1$, the autoregressive coefficient on the GARCH innovations converges to zero, as such the persistence of variance shocks will be very limited.

4.2. Comparison of Impulse Response Functions

The inclusion of GARCH effects in the calculation of the impulse response function results in strikingly different adjustment paths to monetary policy and technology shocks. In all cases such effects are far more pronounced and the approach to equilibrium is far slower than otherwise. Figures 6 and 7 plot the impulse response functions for our two cases respectively.

In the case of high ARCH and low persistent the adjustment paths are fairly similar to

Figure 7: Case 2: High ARCH low persistence $a_z = a_r = 0.8$



the case where such effects are suppressed, albeit with a slightly higher impact and longer delay in approaching the long-run level. However, in the case of low ARCH and high persistence the system exhibits substantial 'excess' volatility as the adjustment to equilibrium is not monotonic. The system adjusts to its long-run equilibrium in a cyclical manner following a shock with the highest impact on the state variables materialising some periods after the materialisation of the shock. Ignoring such variance effects will lead to a serious miscalculation regarding the effect of monetary policy on output, inflation and consumption. They will underestimate the effect on output and the path of inflation. Importantly, inflation will reach a peak after the monetary policy shock has taken effect and it will take substantially longer to work through the system, see Figures 6 subplot (3,3). In addition, the effect on consumption reaches its minimum after the shock although initially, consumption appears less effected by monetary policy.

5. Concluding Remarks

We have formulated a methodology for constructing discrete impulse response functions in mean, variance and covariance for multi-equation models that exhibit autoregressive conditional heteroscedasticity in the model disturbances. This model has many attractive properties and treats the modelling of the first and second moments in a unified and comprehensive manner. When the MV-GARCH effects are ignored the usually reported impulse response functions are distortions of the true adjustment path. In the presence of MV-GARCH, unadjusted IRFs will not appropriately represent both the magnitude of the disturbance to the system and the number of periods required for return to the long run equilibrium. Using quarterly macroeconomic models for the US economy we find substantial evidence of MV-GARCH effects and demonstrated that IRFs computed without them ignore a great deal of information both in terms of evolution and scale to the major macroeconomic variables following shocks. In the context of this structural VAR model our results suggest that in ignoring MV-GARCH effects, policy responses will be misdirected in achieving reductions in unemployment following negative demand shocks and will underestimate the impact of supply shocks to output. The widely used DSGE models used extensively for policy evaluation will tend to show smooth and relatively rapid adjustments following policy innovations,given their chosen calibrated parameters, but if they ignore potential GARCH effects such time profiles will prove unnecessarily optimistic and the emerging real response will frustrate policy makers whose expectations were based on the incomplete understanding of the impact of the second moments on the system's responses.

References

An, S. and F. Schorfheide (2007). Bayesian analysis of dsge models. *Econometric reviews 26*(2-4), 113–172.

Asker, J., A. Collard-Wexler, and J. De-Loecker (2012). Productivity volatility and the misallocation of resources in developing economies. Technical Report wp17175, NBER.

Benigno, G., P. Benigno, and S. Nisticò (2011). Risk, monetary policy and the exchange rate. Technical report, National Bureau of Economic Research.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review 79*(4), 655–73.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica 77*(3), 623–685.

Bullard, J. and J. W. Keating (1995, December). The long-run relationship between inflation and output in postwar economies. *Journal of Monetary Economics 36*(3), 477–496.

DeJong, D. N., R. Liesenfeld, and J. F. Richard (2006). Timing structural change: a conditional probabilistic approach. *Journal of Applied Econometrics 21*(2), 175–190.

Ding, Z. and R. Engle (2001). Large scale conditional covariance matrix modeling, estimation and testing. *Academia Economic Papers 29*(2), 157–184.

Elder, J. (2003). An impulse response function for a vector autoregression with multivariate GARCH in mean. *Economics Letters 79*(1), 21–26.

Engle, R., D. Lilien, and R. Robbins (1987). Estimating time varying risk premia in the term structure: The arch-m model. *Econometrica 55*, 391–407.

Engle, R. F. and F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory 11*(1), 122–150.

Engle, R. F. and J. G. Rangel (2008). The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies 21*(3), 1187–1222.

Gallant, A. R., P. E. Rossi, and G. Tauchen (1993). Nonlinear dynamic structures. *Econometrica 61*(4), 871–907.

Gourieroux, C. and J. Jasiak (1999). *Financial Econometrics*. Princeton University Press.

Grey, S. (1996). Modelling the conditional distribution of interest rates as a regime switching process. *Journal of Financial Economics 42*, 27–62.

Hafner, C. and H. Herwartz (2001). Volatility impulse response functions for multivariate GARCH models. Technical report, CORE Discussion Paper 2001/3.

Jorda, O. (2005). Estimation and inference of impulse responses by local projections. *The American Economic Review 95(1)*, 161–182.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2011). Investment shocks and the relative price of investment. *Review of Economic Dynamics 14*(1), 102–121.

King, R. G. and M. W. Watson (1994). The post-war u.s. phillips curve: a revisionist econometric history. *Carnegie-Rochester Conference Series on Public Policy 41*(1), 157–219.

Komunjer, I. and S. Ng (2011). Dynamic identification of dynamic stochastic general equilibrium models. *Econometrica 79*(6), 1995–2032.

Koop, G., M. N. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*(1), 119–147.

Lee, J. (2006). The comovement between output and prices: Evidence from a dynamic conditional correlation GARCH model. *Economics Letters 91*(1), 110–116.

Lin, W. (1997). Impulse response function for conditional volatility in GARCH models. *Journal of Business and Economic Statistics 15*(1), 15–25.

Morris, S. D. (2012a). Analytical solutions of dsge models and instrumental variables. Working paper, University of California, San Diego working paper.

Morris, S. D. (2012b). Var (1) representation of dsge models. Technical report, University of California, San Diego working paper.

Pancrazi, R. and M. Vukotic (2011). Tfp persistence and monetary policy. Technical report, Working paper Tolouse School of Economics.

Polasek, W. and L. Ren (1998). Model selection for VAR-GARCH models using the BASEL package. Technical report, Institute of Statistics and Econometrics, University of Basel.

Polasek, W. and L. Ren (2000). Generalised impulse response functions for VAR-GARCH-M models. Technical report, Institute of Statistics and Econometrics, University of Basel.

Sin, C. Y. (2005). Impulse response and variance decomposition in the presence of time varying volatility: with applications to price discovery. Technical report, Mimeo, Department of Economics, Hong Kong Baptist University.

A. Stability and Alternative MV-GARCH Specifications

Two common specifications of the covariance process are the MARCH model of Ding and Engle (2001) and the BEKK model of Engle and Kroner (1995). We now review both models in some detail for our purposes. Setting:

$$-A(L)\,y_t = c + C(L)\,u_t \tag{58}$$

$$-S(L)\,\Sigma_t = QQ' + G(L)\,u_t u'_t \tag{59}$$

where $u_t = y_t - \mathbb{E}y_{t|t-1,\ldots,t_0}$ and $S(L)\,\Sigma_t = \Sigma_t + f^1\Sigma_{t-1} + f^2\Sigma_{t-1} + \ldots + f^q\Sigma_{t-1}$ and $G(L)\,u_t u'_t = g^1 u_{t-1}u'_{t-1} + g^2 u_{t-2}u'_{t-2} + g^3 u_{t-3}u'_{t-3} + \ldots + g^1 u_{t-p}u'_{t-p}$. Here $f^i, \forall i \in 1,\ldots,q$ and $g^i, \forall i \in 1,\ldots,p$ are arbitrary matrix functions such that $f^i : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ and $g^i : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ where $\mathbb{C}^{n\times n}$ is the set of all $n \times n$ positive semi-definite matrices.

A.1. The MARCH Model

The ethos of the MARCH model is to design a simple approach to modelling high-variate conditional covariance systems. The model offers a relatively simple set of independent dynamics and despite this simplicity, it is still computationally very intensive to estimate for larger systems. In the compact notation $f^i\Sigma_{t-i} = B_i \circ \Sigma_{t-i}$ and $g^i\Sigma_{t-i} = A_i \circ u_{t-i}u'_{t-i}$ as such the system variance/covariance matrix is:

$$\Sigma_t = QQ' + \sum_{i=1}^{p} A_i \circ \left(u_{t-i}u'_{t-i}\right) + \sum_{j=1}^{q} B_j \circ \Sigma_{t-j} \tag{60}$$

where $A_i \in \mathbb{C}^{n\times n}$, $B_i \in \mathbb{C}^{n\times n}$ and $\bar{\Sigma} \in \mathbb{C}^{n\times n}$ are parameter matrices. Using the *ivech* transformation

$$A_i = (ivecha_i)\,(ivecha_i)' \tag{61}$$

$$B_j = (ivechb_j)\,(ivechb_j)' \tag{62}$$

$$QQ' = (ivechq)\,(ivechq)' \tag{63}$$

The individual parameter vectors are now $q \in \mathbb{R}^{n(n+1)\times 1}$, $a_i \in \mathbb{R}^{n(n+1)\times 1}$ and $b_j \in \mathbb{R}^{n(n+1)\times 1}$ and the full specification parameter θ vector is therefore,

$$\theta = \left[q', a'_1, ..., a'_p, b'_1, ..., b'_q\right]' \tag{64}$$

The stationarity conditions require

$$\mathbb{P}\left(tr\left(\Sigma_{t|t\in[1,...,\tau]}\right) < \infty\right) = 1 \tag{65}$$

given the following

$$A_i = [a_{i,j,k}] \tag{66}$$
$$B_i = [b_{i,j,k}] \tag{67}$$

and

$$\mathcal{A} = [A_1, ..., A_p] \tag{68}$$
$$\mathcal{B} = [B_1, ..., B_q] \tag{69}$$

permutation and extraction yields

$$\underset{p\times 1}{\lambda_{j,k}} = [a_{1,j,k}, a_{2,j,k}, ..., a_{p,j,k}]' \tag{70}$$
$$\underset{q\times 1}{\gamma_{j,k}} = [b_{1,j,k}, b_{2,j,k}, ..., b_{q,j,k}]' \tag{71}$$

For stationarity the roots of the following sets of polynomial pairs

$$1 - \left(\lambda_{1,j,k}z + \lambda_{2,j,k}z^2 + ... + \lambda_{p,j,k}z^p\right) = 0 \tag{72}$$
$$1 - \left(\gamma_{1,j,k}z + \gamma_{2,j,k}z^2 + ... + \gamma_{p,j,k}z^p\right) = 0 \tag{73}$$

must lie in the [0,1] domain for stationarity and non-negativity. By setting $a_i \in \mathbb{R}_+$ and $b_i \in \mathbb{R}_+$, this simplifies to, $\sum_{i=1}^{p} \lambda_i = 1$ and $\sum_{i=1}^{q} \gamma_i = 1$. The MARCH specification is typically used in the construction of very large conditional covariance matrices. However even with the use of the simple element by element approach to the time varying dynamics maximum likelihood estimation still requires the construction and inversion of a large number of non-negative matrices.

A.2. The BEKK Model

An alternative specification to the MARCH model is the BEKK specification of Engle and Kroner (1995), $f^i\Sigma_{t-i} = B_i'\Sigma_{t-i}B_i$ and $g^i\Sigma_{t-i} = A_i'u_{t-i}u_{t-i}'A_i$. This appears to be an attractive feature when considering fairly large scale covariance matrices. In particular, the specification structure allows the BEKK to capture localized effects in the off-diagonal elements and still ensures that Σ_t is non-negative.

$$\Sigma_t = QQ' + \sum_{i=1}^{p} A_i' \left(u_{t-i} u_{t-i}^{'} \right) A_i + \sum_{j=1}^{q} B_i^{'} \Sigma_{t-j} B_i \tag{74}$$

An attractive feature of the BEKK model is the simple partitioned structure of the parameter vector. Only the parameters forming the unconditional covariance matrix $\bar{\Sigma}$, need to be parsed via the *vech* transformation to ensure that Σ_t is non-negative definite. As such if $A_i \in \mathbb{R}^{n \times n}$ and $B_i \in \mathbb{R}^{n \times n}$, the parameter vector is:

$$\theta = \left[q', vecA_1^{'}, ..., vecA_p', vecB_1^{'}, ..., vecB_q' \right]' \tag{75}$$

with $\bar{\Sigma}$ the same as in the MARCH specification. Both models treat the recursion of the covariance matrix as a matrix autoregressive moving average process. Assuming that the conditional distribution is multivariate normal $\mathfrak{F}_t (u_t)$ the time t log-likelihood score is therefore,

$$\log \mathfrak{F}_t \left(u_t \left| \theta \right. \right) = -\tfrac{1}{2} n \log 2\pi - \tfrac{1}{2} \log |\Sigma_t| - \tfrac{1}{2} u_t' \Sigma_t^{-1} u_t \tag{76}$$

the overall objective log-likelihood function $\mathfrak{L}(\theta)$ is:

$$\mathfrak{L}(\theta) = -\tfrac{1}{2} \tau n \log 2\pi - \tfrac{1}{2} \sum_{t=1}^{\tau} \log |\Sigma_t| + u_t' \Sigma_t^{-1} u_t \tag{77}$$

Estimation of the parameters proceeds using a standard non-linear optimization approach. The stationarity constraints on the BEKK model are very simple primarily due to the quadratic form of the model. If $\mathcal{A} = [A_1, ..., A_p]$ and $\mathcal{B} = [B_1, ..., B_q]$ are the arrays of parameter matrices for a particular model then the unconditional (long run) covariance Σ matrix is:

$$vec(\Sigma) = \left(I - \sum_{i=1}^{p} (A_j \otimes A_j)^{'} - \sum_{i=1}^{q} (B_j \otimes B_j)' \right)^{-1} vec(QQ') \tag{78}$$

where I is an $\frac{1}{2}n(n+1)$ dimension identity matrix.