ARTICLE IN PRESS

International Journal of Forecasting ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting with VAR models: Fat tails and stochastic volatility

Ching-Wai (Jeremy) Chiu[a], Haroon Mumtaz[b,*], Gábor Pintér[a]

[a] Bank of England, Threadneedle Street, London, EC2R 8AH, United Kingdom

[b] School of Economics and Finance, Queen Mary University of London, London, E1 4NS, United Kingdom

ARTICLE INFO

Keywords:
Bayesian VAR
Fat-tails
Stochastic volatility
Great Recession

ABSTRACT

We provide evidence that modelling both fat tails and stochastic volatility are important in improving in-sample fit and out-of-sample forecasting performance. To show this, we construct a VAR model where the orthogonalised shocks feature Student's t distribution as well as time-varying variance. We estimate the model using US data on industrial production growth, inflation, interest rates and stock returns. In terms of in-sample fit, the VAR model featuring both stochastic volatility and t-distributed disturbances outperforms restricted alternatives that feature either attributes. The VAR model with t disturbances results in density forecasts for industrial production and stock returns that are superior to alternatives that assume Gaussianity, and this difference is especially stark over the recent Great Recession. Further international evidence confirms that accounting for both stochastic volatility and Student's t-distributed disturbances may lead to improved forecast accuracy.

1. Introduction

Could empirical macroeconomic models with a more realistic shock distribution be able to better predict economic downturns? Since the Great Recession and during the ensuing uncertainty surrounding the political and economic environment, both academic and policy circles have paid increasing attention to fat tail events. Many argue that recent events could hardly be explained or predicted by models that are based on a Gaussian shock structure, mainly because these models assign virtually zero probability to the macroeconomic outcomes that we have recently observed.[1] This has been recognised by recent efforts of the DSGE literature including Chib and Ramamurthy (2014) and Curdia, del Negro, and Greenwald (2014) who found evidence that models with a multivariate t-distributed shock structure are strongly favoured by the data over standard Gaussian models.

This paper contributes to the literature by empirically investigating the in-sample fit and out-of-sample forecasting performance of a VAR model incorporated with Student's t errors (Student, 1908) and stochastic volatility (TVARSVOL). Building on the previous work on univariate (Geweke, 1992, 1993, 1994) and multivariate (Ni & Sun, 2005) models with Student's t-distributed shocks, as well as work on the DSGE literature (Fernandez-Villaverde & Rubio-Ramirez, 2007; Justiniano & Primiceri, 2008; Liu, Waggoner, & Zha, 2011) on stochastic volatility of the error structure, we provide a Gibbs sampling algorithm to estimate the TVARSVOL model. Moreover, we apply the particle filter to compute the marginal likelihood, and compare the in-sample fit and the out-of-sample forecasting performance of this model against three other models, namely,

* Corresponding author.
E-mail addresses: jeremy.chiu@bankofengland.co.uk (C.-W. Chiu), h.mumtaz@qmul.ac.uk (H. Mumtaz), gabor.pinter@bankofengland.co.uk (G. Pinter).

[1] These issues have been discussed in more detail by Elliott and Timmermann (2013), Mishkin (2011) and Ng and Wright (2013) amongst many others.

http://dx.doi.org/10.1016/j.ijforecast.2017.03.001
0169-2070/

Please cite this article in press as: Chiu, C.-W., et al., Forecasting with VAR models: Fat tails and stochastic volatility. International Journal of Forecasting (2017), http://dx.doi.org/10.1016/j.ijforecast.2017.03.001

a linear Gaussian BVAR model (BVAR), a VAR model with Student's t error (TVAR) and a VAR model with stochastic volatility (VARSVOL).

We show that incorporating both fat tails and stochastic volatility can be important in improving in-sample fit and out-of-sample forecasting performance. Using monthly data on industrial production growth, inflation rate, short-term interest rate and the SP500 return for the US, the TVARSVOL model outperforms the other three models in terms of in-sample fit. When it comes to out-of-sample forecasting, we present international evidence that VAR models with Student's t-distributed shocks result in density forecasts for industrial production and stock returns being superior to alternatives that assume Normality.

Our results have at least two important implications when interpreting historical data. First, the structural shift in industrial production volatility in the early 1980s, often referred to as the Great Moderation, may be over-estimated when the VAR model does not account for Student's t-disturbances. Second, the Student's t assumption appears especially important over the 2008 and 2009 period. Forecast densities for industrial production generated from VARs with Gaussian disturbances assign a negligible probability to the collapse of industrial production observed in late 2008. In contrast, when Student's t shocks are incorporated, the left tail of the forecast density includes the actual outcome.

Our paper is related to the DSGE analysis of Curdia et al. (2014) who show that by solely focusing on fat-tails and ignoring lower-frequency changes in the volatility of shocks (as in Ascari, Fagiolo, & Roventini, forthcoming) tends to bias the results towards finding evidence in favour of fat tails. Our work is also related to Clark and Ravazzolo (forthcoming) who work with (V)AR models using quarterly real-time data (GDP growth rate, inflation rate, unemployment rate and short-term interest rate) of the US. They find empirical evidence that models with stochastic volatility increase the accuracy of both point and density forecasts relative to models assuming homoscedasticity. Our paper considers monthly data-sets incorporating with both real and financial variables from the US and three other developed countries, and provide evidence that modelling fat-tailed errors on top of stochastic volatility is important in improving forecasting performance.

The structure of the paper is as follows. Section 2 provides a description of the TVARSVOL model together with the priors and the conditional posteriors and the computation of the marginal likelihood. This section also describes the restricted models considered in our study. Section 3 presents the posterior estimates, compares the models based on in-sample fit and forecasting performance, and provides sensitivity analysis. Section 4 provides further international evidence on the forecasting performance of the different models estimated on data from Canada, Germany and the UK. Section 5 concludes.

2. The model

The model presented in this section is a multivariate time series model with both time varying variance covariance matrix and Student-t distributed shocks in each of the equations (denoted by TVARSVOL). Stochastic volatility is meant to capture possible heteroscedasticity of the shocks and potential nonlinearities in the dynamic relationships of the model variables, which are related to the low-frequency changes in the volatility.[2] Introducing Student's t-distribution in the shock structure is meant to capture high-frequency changes in volatility that are often of extreme magnitudes, hence potentially providing an effective treatment of outliers and extreme events.[3] By allowing for stochastic volatility and t-distributed shocks, we let the data determine whether time variation in the model structure derives from rare but potentially transient events, or from persistent shifts in the volatility regime.

Consider a simple VAR model:

$$Y_t = c + B_1 y_{t-1} + \cdots + B_p y_{t-p} + u_t \quad t = 1, \ldots, T, \tag{2.1}$$

where y_t is an $n \times 1$ vector of observed endogenous variables, and c is an $n \times 1$ vector of constants; B_i, $i = 1, \ldots, p$ are $n \times n$ matrices of coefficients; u_t are heteroscedastic shocks associated with the VAR equations. In particular, we assume that the covariance matrix of u_t is defined as:

$$cov(u_t) = \Sigma_t = A^{-1} H_t A^{-1'}, \tag{2.2}$$

where A is a lower triangular matrix and $H_t = diag\left(\sigma^2_{1,t}\frac{1}{\lambda_{1,t}}, \sigma^2_{2,t}\frac{1}{\lambda_{2,t}}, \ldots, \sigma^2_{n,t}\frac{1}{\lambda_{n,t}}\right)$ with:

$$\ln \sigma_{k,t} = \ln \sigma_{k,t-1} + s_{k,t}, \; var(s_{k,t}) = g_k, \tag{2.3}$$

for $k = 1, 2, \ldots, n$. In line with Geweke (1993), the weights $[\lambda_{1,t}, \lambda_{2,t}, \ldots, \lambda_{n,t}]$ are indexed by time t because they are to capture any high-frequency movements in volatility over time, as opposed to the low-frequency movements in volatility which are in turn captured by $[\sigma_{1,t}, \sigma_{2,t}, \ldots, \sigma_{n,t}]$. As shown by Geweke (1993), assuming a Gamma prior for $\lambda_{k,t}$ of the form $p(\lambda_k) = \prod_{t=1}^{T} p(\lambda_{k,t}) = \prod_{t=1}^{T} \tilde{\Gamma}(1, \upsilon_{\lambda,k})$ leads to a scale mixture of normals for the orthogonal residuals $\tilde{\varepsilon}_t = Au_t$ where $\tilde{\varepsilon}_t = \{\tilde{\varepsilon}_{1,t}, \tilde{\varepsilon}_{2,t}, \ldots \tilde{\varepsilon}_{n,t}\}$ and $cov(\tilde{\varepsilon}_t) = H_t$. Note that $\tilde{\Gamma}(a, b)$ denotes a Gamma density with mean a and degrees of freedom b.[4] The above formulation is equivalent to a specification that assumes Student's t-distribution for $\tilde{\varepsilon}_{k,t}$ with $\upsilon_{\lambda,k}$ degrees of freedom. Our specification allows the variance of this density to change over time via Eq. (2.3).

There are two noteworthy things about the BVAR model. First, as discussed earlier, it allows for both low and high

[2] See Primiceri (2005) and Uhlig (1997) amongst many others.

[3] In an important paper, Jacquier, Polson, and Rossi (2004) provides a detailed analysis of this issue in a univariate framework. Also see Gerlach, Carter, and Kohn (2000) for a rich discussion of outliers in a Bayesian context, and Bauwens, Koop, Korobilis, and Rombouts (2015) for exploring the role of structural changes in affecting Bayesian forecasts.

[4] The probability density function for the Gamma distribution is:

$$\tilde{\Gamma}(\mu, \upsilon) = \begin{cases} \left[\left(\frac{2\mu}{\upsilon}\right)^{\upsilon/2} \Gamma\left(\frac{\upsilon}{2}\right)\right]^{-1} y^{\frac{\upsilon-2}{2}} \exp\left(-\frac{y\upsilon}{2\mu}\right) & \text{if } 0 < y < \infty \\ 0 \quad \text{otherwise.} \end{cases}$$

Under this parametrisation, the $\tilde{\Gamma}(\upsilon_0, 2)$ distribution is an exponential distribution with mean υ_0.