Forecasting energy price volatilities and comovements with fractionally integrated MGARCH models

Previsione della volatilità e co-movimenti dei prezzi dei prodotti energetici mediante modelli MGARCH frazionalmente integrati

Malvina Marchese and Francesca Di Iorio

Abstract We investigate the use of fractionally integrated MGARCH models from a forecasting and a risk management perspective for energy prices. Our in-sample results show significant evidence of long memory decay in energy price returns volatilities, of leverage effects and of time-varying autocorrelations. The forecasting performance of the models is assessed by the SPA test, the Model Confidence Set and the Value at Risk.
Abstract *I modelli MGARCH frazionalmente integrati vengono impiegti in questo lavoro per lo previsione e la gestione del rischio nel mercato dei prodotti energetici. I risultati ottenuti mostrano la presenza di long-memory nella volatilità dei rendimenti dei prezzi, effetti di leverage e autocorrelatione variante nel tempo. La capacit previsiva di questi modelli è stata verificata mediante il test SPA, il Model Confidence Set e il Value at Risk.*

Key words: Multivariate GARCH models, Long memory, SPA test, MCS, VaR

1 Introduction

The effects of oil price shocks on macroeconomic variables, the impact on the financial sector of energy prices fluctuations, the degree of integration between different energy markets, have been the main fields of a increasing studies. It seems to be a growing consensus in the literature on the use of multivariate GARCH (MGARCH) models. A main reference can be Wang and Wu (2012) that compare the forecasting performances of several univariate and bivariate GARCH-type models for spot

Malvina Marchese
Cass Business School, 106 Bunhill Row London EC1Y 8TZ, e-mail: malvina.marchese@city.ac.uk

Francesca Di Iorio
Università di Napoli Federico II, via L. Rodinò 22 80138 Napoli, Italia e-mail: fdiiorio@unina.it

price returns of crude oil (WTI), conventional gasoline (NYH), heating oil (NYH) and jet fuel. Using several univariate loss functions, they find that the full and diagonal BEKK cannot be outperformed by any other model according to the superior predictive ability (SPA) test of Hansen (2005). Growing attention has been devoted to the problem of an accurate modelling and forecasting of energy price volatilities and correlations for portfolio construction, hedging purposes and risk management. The multivariate model used to these ends should capture all the statistical regularities of the volatility series and allow for time dependent correlations. While there is significant empirical evidence that univariate energy price volatilities display a strong degree of persistence, consistent with a long memory structure (e.g. Chang et al., 2010), it would seem that no attempt to include it in multivariate models has yet been made. More, choosing the most appropriate MGARCH specification must also entail comparison of the models' forecasting abilities and usefulness in a decision based framework. Most forecasting comparisons in the energy literature are based on univariate loss functions and univariate tests such as the Diebold-Mariano, that do not allow for joint evaluation of volatilities and correlations forecasts accuracy. Indeed a comprehensive conditional variance matrix forecasts comparison based on matrix loss functions seems to be lacking so far. This paper investigates and analyze the comovements across three major energy markets, namely crude oil (West Texas Intermediate-Cushing Oklaoma), conventional gasoline (New York Harbor) and heating oil (New York Harbor), by means of several multivariate GARCH-type models with particular attention on the fractionally integrated dynamic conditional correlation (FI-DCC) model. This multivariate GARCH models with long memory, asymmetries and dynamic correlations significantly improves the models' in sample and forecasting performance, and then the attractiveness in terms of risk monitoring of this class of models.

The aims of this papers are: (i) compare the in sample performances of several alternative multivariate GARCH models for the returns on the spot prices of the three series using standard information criteria; (ii) evaluate their forecasting accuracy using the Superior Predictive Ability (SPA) test of Hansen (2005) and the Model Confidence Set (MCS) method of Hansen, Lunde and Nason (2011); (iii) explore the efficiency gains in using the fractionally integrated DCC model for one step ahead Value at Risk prediction for short and long positions.

2 Data and models estimations

The series under investigation are the spot energy price returns of the crude oil (CO), conventional gasoline (CG) and heating oil (HO). We obtain 6401 valid daily observations from June 1st 1992 till June 12th 2017 from the Energy Information Administration (EIA) in the U.S. Department of Energy. Each series show the expected stylized facts, such as non normality and fat tails, and the Ljung and Box Q shows that the null hypothesis of no autocorrelation up to the 10th lag is rejected at 10% level of significance. The presence of long memory for the returns $\boldsymbol{r}_t$ is assessed by the significant estimation of the long-memory parameter d using the local Whittle estimator. To account for the serial correlation found in the data, we fit a VAR model to the returns vector whit 1 Lag as suggested by selection criteria. Only conventional gasoline displays time dependence in the mean equation and there are no evidences of spillover effects between the series. Post-estimation diagnostic tests for the residuals of the estimated VAR(1) model confirm the presence of strong GARCH effects, non Normality and no serial correlation up to lag 20. Based on the above evidences we fit the several multivariate GARCH specifications to the VAR(1) residuals. Denoting by $\boldsymbol{r}_t$ the vector of log- returns of n oil prices and θ a finite vector of parameters, the general form of a multivariate GARCH (MGARCH) model is: $\boldsymbol{r}_t = \boldsymbol{\mu}_t(\theta) + \boldsymbol{\varepsilon}_t$, with $\boldsymbol{\varepsilon}_t = \boldsymbol{H}_t^{1/2}(\theta)\boldsymbol{z}_t$ where $\boldsymbol{z}_t$ is an *i.i.d* zero mean random vector such that $Var(\boldsymbol{z}_t) = \boldsymbol{I}_n$, and $\boldsymbol{H}_t^{1/2}$ is a $n \times n$ positive definite matrix. The conditional mean of the process is $\boldsymbol{\mu}_t(\theta)$, the matrix $\boldsymbol{H}_t(\theta) = \boldsymbol{H}_t^{1/2}\boldsymbol{I}_n\left(\boldsymbol{H}_t^{1/2}\right)'$ is the conditional variance. The MGARCH specifications are based on different parameterizations of $\boldsymbol{H}_t$ which have been proposed to capture the dynamics of volatilities and correlations, avoiding the *curse of dimensionality* and to ensure positive definiteness of the covariance matrix. Comprehensive reviews of multivariate GARCH models can be found in Bauwens *et al* (2006) and Silvennoinen and Terasvirta (2009). The MGARCH models estimated in the paper and their main characteristics are summarized in Table 1. In this paper the model estimation is performed by Maximum Likelihood methods in one step under the assumption of joint normality of the vector of disturbances using Kevin Sheppard's MFE Toolbox for *Matlab*, release 2016a. Standardized residuals of estimated volatility models are fat tailed, so the assumption of Gaussianity of the innovations is not innocuous and reduces efficiency; however Gaussian likelihood retains consistency under misspecification of the conditional density, as long as the conditional mean and the conditional variance are correctly specified. We estimate $\hat{\theta}$ by maximizing the conditional log likelihood: $L_T(\theta) = c - \frac{1}{2}\sum_{t=1}^{T}\ln|\boldsymbol{H}_T| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{r}_t'\boldsymbol{H}_t^{-1}\boldsymbol{r}_t$.

Table 1 MGARCH models and their characteristics

Models	Dynamic Corr.	Asymmetries	L-Memory	Spillovers
DBEKK	x			
BEKK	x			x
ABEKK	x	x		x
AGARCH		x		x
CCC				
DCC	x			
cDCC	x			
FI-DCC	x		x	
FI-EDCC		x	x	

Full results concerning the model estimations, obtained with one-step Maximum Likelihood, are available from the authors on request. Table 2 reports the maximized log-likelihood and information criteria for all the fitted models. 'Np' is the number of estimated parameters in each model and the values in bold correspond to the best performing models. Weather the criterion is AIC or BIC, constant conditional correlation specifications are outperformed by their dynamic counterparts and symmetric specifications are outperformed by specifications including leverage effects.

Table 2 Maximized log-likelihood and information criteria

Models	Np	LogLik	AIC	BIC
DBEKK	12	-18321	36666	36737
BEKK	24	-18218	36484	36615
ABEKK	33	-17989	36044	36240
AGARCH	33	-18011	36088	36284
CCC	12	-18202	36428	36499
DCC	14	-17695	35414	**35501**
FI-DCC	17	-17684	**35402**	**35503**
FI-EDCC	23	-17661	**35368**	35505

3 The forecasting exercise

Evaluation of volatility forecasts is particularly challenging since volatility itself is latent and thus unobservable even ex post. In this case to compare model based forecasts with ex post realizations a statistical or an economic loss function as well as a proxy for the true unobservable conditional variance matrix have to be chosen. Proxy might lead to a different ordering of competing models that would be obtained if the true volatility were observed. To avoid a distorted outcome, the choice of an appropriate loss function is crucial. In this paper, we follow Bauwens et al. (2016) and use several loss functions, robust to noisy proxies. i.e. expected to provide the same forecasts ranking using the true conditional covariance or a conditionally unbiased proxy. Their definition is provided in Table 3 where H_{it} , for $i = 1,\dots,k$ denotes each model predicted covariance matrix for day t, $\hat{\Sigma}_t$ is the proxy of the conditional covariance matrix, ι is a vector of ones, T is the out of sample length and n is the the sample size. As a proxy for the conditional variance matrix at day t we use the matrix of the outer products of the daily mean forecast errors, $e_{T+1}e'_{T+1}$ which is a conditionally unbiased proxy. The forecasting ability of the set of proposed models is evaluated over a series of 630 out-of sample predictions. We compare the one day ahead conditional variance matrix forecasts based on the models estimated. We divide the full data set into two periods: the *in-sample period* from 02 August 2004 to 9 January 2014 (2430 observations), the *out of sample* with 510 observations from 10 January 2014 to 31 December 2015, used for forecasting evaluation. Forecasts are constructed using a fixed rolling window scheme: the estimation period is rolled forward by adding one new daily observation and dropping the most distant observation. Models parameters are re-estimated each day to obtain tomorrow volatility forecasts and the sample size employed to estimate is fixed and any dependence on the mean dynamics has been accounted for by fitting a VAR(1), so the mean forecasts do not depend on the models. This scheme satisfies the assumptions required by the MCS method and the SPA test and allows a unified treatment of nested and non-nested models. For each statistical loss function, we evaluate the significance of loss functions differences by means of the SPA and MCS.

Table 3 Loss function

Loss function		Type
Frobenius	$tr\left[\left(\hat{\Sigma}_t - H_{it}\right)'\left(\hat{\Sigma}_t - H_{it}\right)\right]$	Symmetric
Euclidean	$vech\left(\hat{\Sigma}_t - H_{it}\right)' vech\left(\hat{\Sigma}_t - H_{it}\right)$	Symmetric
MSFE	$\frac{1}{T} vec\left(\hat{\Sigma}_t - H_{it}\right)' vec\left(\hat{\Sigma}_t - H_{it}\right)'$	Symmetric
QLIKE	$\log\lvert H_t\rvert + vec\left(H_{it}^{-1}\hat{\Sigma}_t\right)' \iota$	Symmetric
Stein	$tr\left(H_{it}^{-1}\hat{\Sigma}_t\right) - \log\left\lvert H_{it}^{-1}\hat{\Sigma}_t\right\rvert - n$	Asymmetric
VDN	$tr(\hat{\Sigma}_t \log\hat{\Sigma}_t - \hat{\Sigma}_t \log H_{it} - \hat{\Sigma}_t + H_{it})$	Symmetric

We follow Hansen (2005) and obtain the p-values of the test by bootstrap. We implement a block bootstrap with block length equal to 2 and 10000 bootstrap samples. We find that the hypothesis of constant correlation is always rejected, as well as the hypothesis of short memory. The hypothesis of symmetry in the volatility dynamics is rejected in most benchmarks, and allowing for dynamic correlations significantly improves the models' forecasting accuracy. In the overall it appears that the most valid specification in this application is the fractionally integrated exponential DCC model that captures well the dynamics of the variance covariance matrix. The MCS methodology identifies a set of models with equivalent predictive ability which outperform all the other competing models at a given confidence level α with respect to a particular loss function. MCS determines the set of models that at a given confidence level have the best forecasting performance. We use a block bootstrap scheme to obtain the quantiles of the distribution. The block length bootstrap parameter is set equal to 2 and the number of bootstrap sample used is 10000. At the 90% confidence level the asymmetric BEKK and the fractionally integrated exponential DCC are included in the MCS resulting from the Euclidean, Frobenious, MSFE and VDN loss functions. The fractionally integrated DCC is included in the MCS deriving from the Euclidean, Frobenious and MSFE loss functions. The highest number of models (eight) is included for the Euclidean and Frobenious loss functions. The most striking result is the inclusion of the fractionally integrated exponential DCC model in the MCS of four loss functions supporting the hypothesis that the inclusion of long memory, asymmetries and time varying correlations significantly improve forecasting accuracy.

The possible efficiency gains of using long-memory asymmetric MGARCH models over short memory benchmarks for one-step ahead Value at Risk forecasting for equally weighted portfolios. To this end, we focus on the models' ability to predict the tail behavior of the returns rather than obtaining the 'best' volatility model.

We forecast the one day ahead Value at Risk for each of the models under comparison at 5%, 2.5% and 1% levels, and we assess their accuracy using statistical back-testing. We are concerned with both the long and short positions *VaR*. So we focus respectively on the left and right tail of the forecasted distribution of returns and we assess the models joint ability to delivery accurate VaR forecasts for both tails. To asses the accuracy of the *VaRs* obtained by the different models we test weather the failure rate implied by each model is statistically equal to the expected one. A popular back-testing procedure is based on the unconditional coverage test of Kupiec (e.g. Giot and Laurent, 2003). The test is a likelihood ratio test, built under the assumption that VaR violations are independent. Under the null, the test statistic is distributed as a χ^2-distribution with two degrees of freedom. Results for the short memory constant correlation models are homogenous for short and long VaRs, leading in all cases to rejection of the null hypothesis, regardless of the model structure. Models with dynamic conditional correlations perform much better passing all the tests with the occasional rejection for the most extreme quantiles. Models with dynamic conditional correlations and long memory adequately forecast *VaRs* at all levels. In conclusion for equally weighted portfolios, reliable *VaR* forecasts can be obtained under the assumption of conditionally normally standardized portfolio returns, by using DCC-type of models that include long range dependence and asymmetries in the individual volatilities.

As a general finding, fractionally integrated dynamic conditional correlation models display good in-sample fit. Using a fixed rolling window scheme, we assess the one-day ahead forecasting accuracy of the models with the MCS method and the SPA test using several matrix loss functions, robust to the choice of the volatility proxy. Short memory constant correlations models are always rejected in favour of long memory dynamic correlation models, and that the use of the latter significantly improves forecasts accuracy from a statistical as well as a risk management perspective.

References

1. Bauwens, L., Laurent, S., Rombouts, J. V., 2006. Multivariate GARCH models: a survey. Journal of applied econometrics 21, 79-109.
2. Bauwens, L., Braione, M., Storti, G., 2016. Forecasting comparison of long term component dynamic models for realized covariance matrices. Annals of Economics and Statistics/Annales d'Économie et de Statistique, (123/124), 103-134.
3. Chang, C. L., McAleer, M., Tansuchat, R., 2010. Analyzing and forecasting volatility spillovers, asymmetries and hedging in major oil markets. Energy Econ. 32, 1445–1455.
4. Hansen, P.R., 2005. A test for superior predictive ability. Journal of Business Economics and Statistics 23, 365–380.
5. Giot, P., Laurent, S., 2003. Value-at-risk for long and short trading positions. Journal of Applied Econometrics 18, 641-663.
6. Silvennoinen, A., Teräsvirta, T. , 2009. Multivariate GARCH models. Handbook of financial time series, 201-229.
7. Wang, Y., Wu, C, 2010. Forecasting energy market volatility using GARCH models: Can multivariate models beat univariate models?. Energy Economics 34, 2167-2181.