Forecasting Crude Oil and Refined Products Volatilities and Correlations: New Evidence from Fractionally-Integrated Multivariate GARCH Models

Malvina Marchese, Ph.D.
Cass Business School, City University of London, U.K.

Michael Tamvakis, Ph.D.
Cass Business School, City University of London, U.K.

Ioannis Kyriakou, Ph.D.
Cass Business School, City University of London, U.K.

Francesca Di Iorio, Ph.D.
Dipartimento di Scienze Politiche, Università degli Studi di Napoli Federico II, Italy

Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3544242

This paper advocates the use of long-memory multivariate GARCH models to forecast spot return volatilities and correlations for crude oil and related products. The findings show from a risk management perspective that the multivariate models incorporating long-memory features outperform the short-memory counterparts in providing the most accurate Value-at-Risk measures. The paper provides useful insights to non-commercial oil traders and other energy markets agents engaged in hedging and risk management operations.

Introduction

There is a consensus in the empirical literature on the effectiveness of multivariate GARCH (MGARCH) models to forecast volatilities and correlations of crude oil and refined products returns. However, all the MGARCH model specifications used in the literature so far implicitly impose a short-memory decay rate on volatilities and correlations. This is problematic since they have been shown to display a strong degree of persistence, i.e., the impact of shocks to them decays very slowly. Several univariate long-memory models, including the fractionally integrated autoregressive (ARFIMA) model and the fractionally integrated GARCH (FIGARCH) model, have been successfully used to forecast the volatilities of crude oil and refined products returns (Block *et al*., 2015; Tong *et al*., 2013; Chang *et al*., 2010; Borenstein *et al*., 1997) but, to the best of our knowledge, no attempt has yet been made to demonstrate the advantage of incorporating the long-memory feature in multivariate models.

In practice, failure to account for this very slow decay rate in the volatility and correlation processes implies misspecification of the true data generating processes which, in turn, can potentially lead to: (i) biased conclusions about the response of refined products volatility to crude oil price shocks, (ii) inaccurate volatility forecasts and (iii) flawed risk management practices.

This digest article was contributed by Ana-Maria Fuertes, Ph.D., Professor in Finance and Econometrics at Cass Business School, City, University of London (U.K.).

This paper fills a gap in the literature by assessing whether, as regards the out-of-sample prediction of volatilities and co-movements between crude oil and refined products returns, the use of multivariate long-memory GARCH models with long-memory leads to gains in statistical accuracy as well as benefits from a risk management perspective.

The paper models the volatilities and correlations of crude oil returns (West Texas Intermediate-Cushing) and two refined products return series, conventional gasoline (New York Harbor) and heating oil (New York Harbor), by means of different MGARCH models, including the fractionally integrated dynamic conditional correlation (DCC) model. The models are rigorously compared in-sample and from an out-of-sample forecasting perspective to assess whether long-memory specifications with dynamic correlations and asymmetries outperform their short memory counterparts. The models' attractiveness in terms of risk management is assessed by forecasting the Value at Risk.

Relevance of the Research Question

Crude oil prices are central to global economic activity. Crude oil is of limited direct usage as a fuel. It is the range of products yielded by refining crude oil which are consumed either directly (e.g., gasoline and diesel for motor vehicles) or indirectly (e.g., fuel oil to generate electricity, or naphtha as petrochemical feedstock). Because of the need to transform crude oil into refined products, the interaction between upstream producers and downstream consumers is not direct. Prices for refined products can be linked back to those of crude oil through the netback mechanism. Refined product prices should theoretically be linked to the cost of acquiring crude oil (of various qualities and provenances), transporting it (via pipelines or tankers, often from abroad) to the transformation point, storing it, refining it, storing the refined products and distributing these products to a myriad of consumption points, which may be located abroad as well. Such calculations might be feasible if all the relevant information were publicly available and easily accessible. As this is not usually the case, researchers investigate the linkages empirically using models estimated with data for the most commonly traded crude oils and refined products.

The subject of the paper is important because return volatilities and correlations of crude oil and refined products are key inputs to macroeconomic models, option pricing models, investment portfolio construction, and hedging and risk management practices inter alia. These practices are of particular significance to the refining industry, which forms the nexus between crude oil production and final consumption and which is exposed to risks from the supply and demand sides of the marketplace.

Data and Models

The paper estimates 48 different MGARCH models using daily spot price returns on crude oil (CO), conventional gasoline (CG) and heating oil (HO) from 1 June 1993 to 1 June 2018 from the Energy Information Administration (EIA) of the U.S. Department of Energy. The daily return is calculated as the difference in the logarithmic closing price.

Examining the data, it is observed that the average daily returns are very small compared to the sample standard deviations. The returns display some evidence of skewness and excess kurtosis (deviation from normality). More importantly for the present purposes, the correlogram and the Ljung–Box Q statistic for

serial correlation of the squared returns suggests a very strong degree of persistence in all volatility series, consistently with a long-memory decay rate. The latter is confirmed by estimating semi-parametrically for each series the long-memory parameter *d* using the local Whittle estimator of Robinson (1995) with bandwidth m = 100 and no trimming. To account for serial correlation in the data, we fit a VAR(p) model to the returns finding that a VAR(1) parameterization suffices to account for the conditional mean dynamics of the series. There is no evidence of spillover effects between the means series.

Results

In-sample results show strong evidence of GARCH-type dynamics, long-range dependence and leverage effects in the individual volatilities. In terms of the multivariate structure, the data strongly support the hypothesis of dynamic conditional correlations.

The most important finding of the paper is that the use of multivariate GARCH models with a *long-memory* significantly improves the out-of-sample forecasting accuracy of volatilities and correlations from the viewpoint of statistical loss functions and economic loss functions.

Using a fixed rolling window scheme, the authors assess the 1-, 5- and 20-day ahead out-of-sample forecasting accuracy of the models using different statistical approaches and criteria (Laurent *et al*., 2012; Hansen, 2005). Since the processes under study (volatilities) are unobservable/latent, the authors consider various matrix loss functions which are robust to the choice of the volatility proxy. Then they evaluate the models' forecasting performance in an economically meaningful way by using the model forecasts as inputs to obtain Value-at-Risk predictions.

The results suggest that models with a long-memory decay rate surpass the short-memory counterparts from a statistical as well as an economic perspective and their use can significantly improve the assessment of oil market risk. The sensitivity of the results to the sample period under study is examined by considering, in addition to the full sample, three sub-samples. The findings indicate that it is particularly important to incorporate long-memory in the multivariate models when the period to be forecasted is a turbulent (as opposed to tranquil) one. Finally, it is also shown that accounting for long-memory in the modeling tools is particularly important when the forecasting horizon is as lengthy as 21 days ahead.

Conclusions

This paper advances research on the modeling of crude oil markets and the markets of refined products by comparing the return volatility and correlation forecasts obtained from multivariate long-memory GARCH models with those obtained from the simpler short-memory models that have been used thus far in the energy markets literature. The results endorse the former and are particularly important for agents including refiners and oil trading companies who have risk exposures to both the crude and refined sides of the market. Risk managers in such companies may consider the long-memory models advocated in this paper to improve their Value-at-Risk forecasts and risk management practices.

References

Block, A. S., Righi, M. B., Schlender, S. G. and D.A. Coronel, 2015, "Investigating Dynamic Conditional Correlation Between Crude Oil and Fuels in Non-Linear Framework: The Financial and Economic Role of Structural Breaks," *Energy Economics,* Vol. 49, May, pp. 23-32.

Borenstein, S., Cameron, A. C. and R. Gilbert, 1997, "Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes?", *The Quarterly Journal of Economics*, Vol. 112, No. 1, February, pp. 305-339.

Chang, C.-L., McAleer, M. and R. Tansuchat, 2010, "Analyzing and Forecasting Volatility Spillovers, Asymmetries and Hedging in Major Oil Markets," *Energy Economics*, Vol. 32, No. 6, November, pp. 1445-1455.

Hansen, P. R., 2005, "A Test for Superior Predictive Ability," *Journal of Business & Economic Statistics,* Vol. 23, No. 4, October, pp. 365-380.

Laurent, S., Rombouts, J. V. K. and F. Violante, 2012, "On the Forecasting Accuracy of Multivariate GARCH Models," *Journal of Applied Econometrics,* Vol. 27, No. 6, September - October, pp. 934-955.

Robinson, P. M., 1995, "Gaussian Semiparametric Estimation of Long Range Dependence," *The Annals of Statistics*, Vol. 23, No. 5, October, pp. 1630-1661.

Tong, B., Wu, C. and C. Zhou, 2013, "Modeling the Co-movements Between Crude Oil and Refined Petroleum Markets," *Energy Economics,* Vol. 40, November, pp. 882-897.

Keywords

Crude oil and refined products correlations, volatility forecasting, multivariate GARCH with long memory, Superior Predictive Ability test, Value-at-Risk.