International Journal of Forecasting 34 (2018) 622–635

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting crude oil price volatility

Ana María Herrera [a], Liang Hu [b,*], Daniel Pastor [c]

[a] Department of Economics, University of Kentucky, United States
[b] Department of Economics, Wayne State University, United States
[c] Department of Economics and Finance, University of Texas at El Paso, United States

ARTICLE INFO

Keywords:
Crude oil price volatility
GARCH models
Long memory
Markov switching
Volatility forecast
Realized volatility

ABSTRACT

We use high-frequency intra-day realized volatility data to evaluate the relative forecasting performances of various models that are used commonly for forecasting the volatility of crude oil daily spot returns at multiple horizons. These models include the RiskMetrics, GARCH, asymmetric GARCH, fractional integrated GARCH and Markov switching GARCH models. We begin by implementing Carrasco, Hu, and Ploberger's (2014) test for regime switching in the mean and variance of the GARCH(1, 1), and find overwhelming support for regime switching. We then perform a comprehensive out-of-sample forecasting performance evaluation using a battery of tests. We find that, under the MSE and QLIKE loss functions: (i) models with a Student's t innovation are favored over those with a normal innovation; (ii) RiskMetrics and GARCH(1, 1) have good predictive accuracies at short forecast horizons, whereas EGARCH(1, 1) yields the most accurate forecasts at medium horizons; and (iii) the Markov switching GARCH shows a superior predictive accuracy at long horizons. These results are established by computing the equal predictive ability test of Diebold and Mariano (1995) and West (1996) and the model confidence set of Hansen, Lunde, and Nason (2011) over the entire evaluation sample. In addition, a comparison of the MSPE ratios computed using a rolling window suggests that the Markov switching GARCH model is better at predicting the volatility during periods of turmoil.

Published by Elsevier B.V. on behalf of International Institute of Forecasters.

1. Introduction

Over recent months, newspaper headlines such as "Oil prices will be much more volatile in 2017: IEA" (Reuters, January 15, 2017) and "IEA sees risk of volatile oil prices on weak upstream investment" (Bloomberg, September 17, 2017) have been evidence of the concerns voiced by the International Energy Agency regarding the return of high volatility in crude oil markets. This time around, the apprehension regarding a higher volatility seems to have stemmed from the slow pace of investment in new production. Nevertheless, surges in the volatility of the daily West Texas Intermediate (WTI) spot returns were observed around the 1986 oil price collapse, during the Gulf War, following the onset of the 2007–2008 financial crisis, and more recently since the fall in oil prices that started in July 2014 (see Fig. 1). Clearly, periods of heightened volatility in crude oil markets are recurrent, and these headlines manifest the importance of evaluating whether the econometric tools that are available to practitioners are able to generate reliable forecasts of the volatility of crude oil pricces.

Since the "spot oil price volatility reflects the volatility of current as well as future values of [oil] production, consumption and inventory demand" (Pindyck, 2004), it is relevant for various economic agents. Accurate forecasts are key for firms whose business depends heavily on oil prices; for instance, oil companies that need to decide whether to drill a new well (Kellogg, 2014) or to undertake long-term investments in their refining and transportation infrastructure, airline companies who use oil price forecasts to set airfares, and the automobile industry. Second,

* Corresponding author.
E-mail address: lianghu@wayne.edu (L. Hu).

https://doi.org/10.1016/j.ijforecast.2018.04.007
0169-2070/Published by Elsevier B.V. on behalf of International Institute of Forecasters.

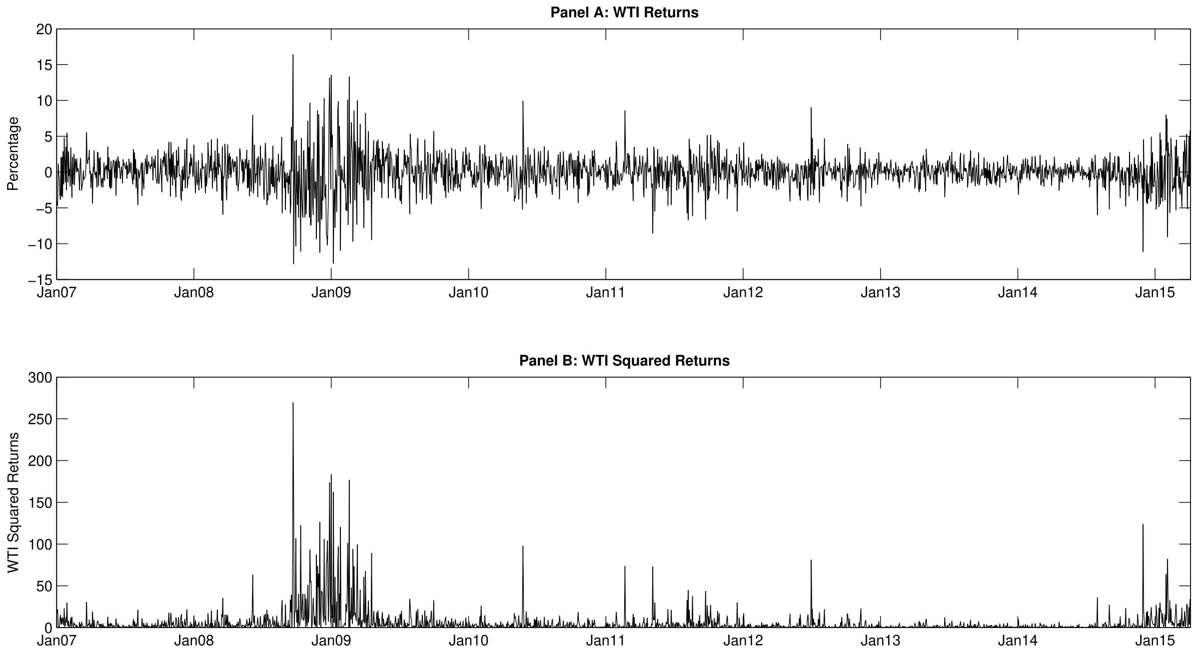

Fig. 1. Daily WTI crude oil returns and squared returns. The sample period extends from January 3, 2007, to April 2, 2015.

the oil price volatility also plays a role in households' decisions regarding their purchases of durable goods (Davis & Kilian, 2011; Kahn, 1986). Lastly, they are useful for agents whose daily task is to produce forecasts of industry-level and aggregate economic activities, such as policy makers, business economists, and private sector forecasters (see e.g. Elder & Serletis, 2010; Jo, 2014).

The aim of this paper is to evaluate the out-of-sample forecasting performances of different volatility models for the conditional variance (hereafter variance) of spot crude oil returns, where we proxy the unobserved variance with the realized volatility of intra-day returns (Andersen & Bollerslev, 1998). More specifically, we investigate the predictive abilities of the RiskMetrics, GARCH, asymmetric GARCH, Fractionally Integrated GARCH (FIGARCH) and Markov switching GARCH (MS-GARCH) models. The motivation for choosing these models is as follows. RiskMetrics remains a very popular empirical model among practitioners. Meanwhile, GARCH (Bollerslev, 1986) sets out the idea of modeling and forecasting the volatility as a time-varying function of currently available information. On the empirical side, the GARCH(1,1) model has also fared well in predicting the conditional volatility of financial assets (Hansen & Lunde, 2005) and the crude oil price volatility (see Xu & Ouenniche, 2012, and references therein). Asymmetric GARCH models such as EGARCH (Nelson, 1991) and GJR-GARCH (Glosten, Jagannathan, & Runkle, 1993) have been shown to have good out-of-sample performances when forecasting the oil price volatility one step ahead (Hou & Suardi, 2012; Mohammadi & Su, 2010). As for Markov switching models, they have been found to be more suitable for modeling situations in which changes in regimes are triggered by sudden shocks to the economy. Thus, they might have good predictive abilities for spot crude oil returns, which are characterized by sudden jumps, due to factors such as political disruptions in the Middle East or military interventions in oil-exporting countries, for instance. However, regime switching and long memory are related intimately, and it is hard to differentiate a Markov switching model from a long memory model (Diebold & Inoue, 2001). Therefore, we add the FIGARCH to our pool of models for forecast evaluation.

We provide a comprehensive study of the relative out-of-sample forecasting performances at multiple horizons. We start by testing formally for regime switches using the procedure proposed by Carrasco, Hu, and Ploberger (2014), then evaluate the directional accuracy using Pesaran and Timmermann's (1992) test. Furthermore, we conduct pairwise comparisons between different candidate models using Diebold and Mariano (1995) and West's (1996) test of equal predictive ability. In addition, we also employ Hansen, Lunde, and Nason's (2011) model confidence set procedure in order to determine the best (set of) model(s) from the pool. All of the tests are reported under two loss functions: the mean square error, MSE, and the quasi likelihood, QLIKE. We also inquire into the stability of the forecasting accuracy of the preferred models over the evaluation period (2013–2014).

Our findings can be summarized as follows: (i) the Student's t distribution is generally favored in the parametric models due to the extremely high kurtosis in the oil return volatility; (ii) the nonparametric model (RiskMetrics) and parsimonious models like GARCH(1,1) perform better at short (1- and 5-day) horizons; (iii) the EGARCH stands out at the 21-day horizon; (iv) the MS-GARCH model yields more accurate forecasts at the longer 63-day horizon; and (v) the MS-GARCH model has a higher predictive ability during periods of turmoil.

We are not the first to consider Markov switching models for forecasting the volatility of the crude oil market. For