Statistics and Probability Letters 119 (2016) 36–44

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

A time-varying long run HEAVY model

Manuela Braione

Université catholique de Louvain, CORE, B-1348 Louvain-La-Neuve, Belgium

ARTICLE INFO

Article history:
Received 23 March 2016
Received in revised form 7 July 2016
Accepted 7 July 2016
Available online 20 July 2016

Keywords:
HEAVY model
DCC
Long term models
Distributed lag
Direct forecasting

ABSTRACT

We propose a scalar variation of the multivariate HEAVY model of Noureldin et al. (2012) featuring a time-varying long run (co)volatility component coupled with DCC dynamics. The new model outperforms the original HEAVY model by delivering more accurate multi-step-ahead predictions.

1. Introduction

There is a vast consensus among practitioners that inclusion of high frequency information enables the development of more accurate forecasting models for the conditional covariance of daily returns. An outstanding example is represented by the class of multivariate High-frEquency-bAsed VolatilitY (HEAVY) models introduced by Noureldin et al. (2012), which links the dynamics of the conditional covariance matrix to the realized measure using a system of two equations akin to the multivariate BEKK (Engle and Kroner, 1995) specification. The model has several advantages, the main ones that it is easy to estimate by MLE and able to provide closed-form forecasting formulas. Nevertheless, when the scalar version of the model is employed with *targeting* (as is often the case in financial applications), the conditional covariance dynamics are driven by only two parameters, thus strongly penalizing the flexibility of the model in times of significantly changing economic conditions. For this reason, the authors raise the interesting question of whether a more sophisticated parameterization could improve the forecasting ability of the model. We address this question by studying a new model, the Time Varying Long Run (TVLR) HEAVY, which extends the baseline HEAVY specification to a component structure that decomposes the conditional covariance matrix into long-run (permanent) and short-run (transitory) components in a multiplicative fashion, similarly to the approach adopted by Golosnoy et al. (2012) and Bauwens et al. (forthcoming). We model the trend component using a finite distributed lag specification typically encountered in the Mixed Data Sampling (MIDAS) regression framework and allow the short term dynamics to move according to a DCC (Engle, 2002) specification, thus stepping away from the basic linear BEKK recursion. We compare the TVLR-HEAVY against the standard HEAVY model and other selected benchmarks from both an in- and out-of-sample perspective. For the TVLR-HEAVY, multi-step ahead forecasts are constructed using the *direct* approach which overcomes the difficulties created by the nonlinear structure of the model. In this way, the model can still be feasibly estimated by MLE, thus keeping computational tractability in practical applications.

Our set of results shows that introducing an additional component that captures the secular movements in the (co)volatility dynamics is well justified, as the new model is found to improve over existing benchmarks both in the overall fit

E-mail address: manuela.braione@uclouvain.be.

http://dx.doi.org/10.1016/j.spl.2016.07.006
0167-7152/

and predictive accuracy. While at very short forecasting horizons other less parameterized models seem to be preferred, the forecast gains of the TVLR-HEAVY tend to be particularly admirable at longer horizons, when the impact of the time-varying trend component appears to be predominant.

The structure of the paper is as follows. Section 2 briefly recalls the multivariate general framework and formally introduces the new model and its estimation approach. Section 3 illustrates the aim of the empirical application and presents the results of both the in- and out-of-sample analysis. Section 4 concludes the paper with some final remarks.

2. General framework

Let $\boldsymbol{r}_t$ denote the $(n \times 1)$ vector of daily returns at time t and $P_t = \boldsymbol{r}_t\boldsymbol{r}_t'$ the $(n \times n)$ matrix obtained as the outer product of daily returns. The realized measure is denoted by V_t, and is a $(n \times n)$, symmetric and positive definite (PD) matrix. Herein we use the realized covariance (RC) estimator obtained by summing up intra-daily returns at the 5 min frequency, although any other consistent estimator could be used.

Conditionally on past information $\Im_{t-1}$ consisting of V_τ for $\tau \leq t-1$, V_t is assumed to follow a n-dimensional central Wishart distribution, i.e. $V_t|\Im_{t-1} \sim W_n(\nu, M_t/\nu)$ with $\nu > (n-1)$, while $P_t|\Im_{t-1} \sim SINGW_n(1, H_t/\nu)$, where $SINGW_n$ denotes a n-dimensional Singular Wishart distribution, following by the assumption that $\boldsymbol{r}_t = H_t^{1/2}\boldsymbol{\epsilon}_t$ with $\boldsymbol{\epsilon}_t \sim N(\boldsymbol{0}, I_n)$. As already stressed in the paper by Noureldin et al. (2012), the distinction between the Wishart and Singular Wishart densities is of no consequence to QML estimation.

Therefore, for the properties of the Wishart distribution, we have that

$$E(P_t|\Im_{t-1}) = E(\boldsymbol{r}_t\boldsymbol{r}_t'|\Im_{t-1}) = H_t \tag{1}$$

$$E(V_t|\Im_{t-1}) = M_t \tag{2}$$

where the PD matrices M_t and H_t are the conditional expectation of the realized measure and of the outer product of daily returns, respectively. Note that they both condition on the same high frequency information, hence they are assumed to be $\Im_{t-1}$ measurable.

The HEAVY model links the dynamics of H_t to the realized measure and is based on a system of two equations for H_t and M_t both akin to the multivariate BEKK specification. Consistently with Noureldin et al. (2012), we will refer to these equations as HEAVY-P and HEAVY-V, unless otherwise stated. Restricting to the scalar case, they are written as follows:

$$H_t = \Omega_H + \alpha_H^2 V_{t-1} + \beta_H^2 H_{t-1} \tag{3}$$

$$M_t = \Omega_M + \alpha_M^2 V_{t-1} + \beta_M^2 M_{t-1} \tag{4}$$

where $\alpha > 0$ and $\beta \geq 0$. If covariance stationarity holds, i.e. $\alpha + \beta < 1$, the model can be expressed in its covariance targeting parameterization. In this case, the intercept matrices Ω_H and Ω_M are written in terms of the unconditional first moments of H_t and M_t and model parameters, i.e. $B_H := E(P_t) = (1-\alpha_H^2-\beta_H^2)^{-1}\Omega_H$ and $B_M := E(V_t) = (1-\alpha_M^2-\beta_M^2)^{-1}\Omega_M$.

Eq. (4) is not needed for computing one-step ahead forecasts of H_t, but is necessary at more than one step ahead due to the presence of V_{t-1} in Eq. (3). Noureldin et al. (2012) propose analytical formulas to achieve multi-period ahead predictions of H_t.

As we will show in a moment, the TVLR-HEAVY model features a nonlinear parameterization of H_t due to the presence of the DCC structure that creates problems in constructing closed-form expressions for multi-step predictions. In a similar scenario, Golosnoy et al. (2012) resorted to indirect forecasting via Monte Carlo simulations of the h-steps ahead forecast distribution of returns. An easier solution, in this framework, can be found by applying the *direct* forecast approach to the HEAVY-P equation, a method that has been extensively used in the empirical finance literature as an alternative to the *iterated* one, see for example Marcellino et al. (2006), Ghysels et al. (2009) and Proietti (2011). It entails to estimate a horizon-specific model of the (co)volatility, say weekly or monthly, which can then be used to form direct predictions over the next week or month. As only observed data are utilized to predict future periods, it is thought to yield reliable results. In this way, only a unidimensional system is needed to achieve direct multi-step ahead predictions of H_t, as those of V_t are directly taken into account in the same equation. We elaborate on this point in the following subsection, which formally introduces the proposed model and its estimation approach.

2.1. The model

The TVLR-HEAVY model features a multiplicative decomposition of the conditional covariance matrix of returns H_t into a secular component $S_t = G_tG_t'$ and a short term component $H_t^\star$, as follows:

$$H_t = G_t H_t^\star G_t' \tag{5}$$

where $H_t^\star$ is a $(n \times n)$ PD matrix and G_t a lower triangular matrix obtained as *e.g.* a Cholesky factorization of S_t. S_t captures the long term movements in the levels around which (co)volatilities fluctuate from day to day while $H_t^\star$ represents the transitory component of the covariance dynamics. In order to identify the model, we impose $E(H_t^\star) = I_n$, with I_n the $(n \times n)$ identity