

Long term component dynamic models for realized covariance matrices

Modelli dinamici a componenti per matrici di covarianza realizzate

Luc Bauwens, Manuela Braione and Giuseppe Storti

Abstract Dynamic models for realized covariance matrices are proposed, which include a secular component that captures the changing levels of realized variances and correlations. They generalize the realized DCC models of Bauwens et al. (2012), where the long term level is assumed to be constant. The long term component is specified either as a nonparametric function or as a MIDAS term. Estimation can be done in steps for large dimensional matrices.

Abstract *Vengono proposti dei modelli dinamici per matrici di covarianza realizzate, che includono una componente capace di catturare i movimenti di lungo periodo delle varianze e correlazioni realizzate. Tali modelli generalizzano i modelli DCC realizzati di Bauwens et al. (2012), dove il livello di lungo periodo di tali variabili si assume essere costante. La componente di lungo periodo può essere specificata in maniera non parametrica o usando una struttura MIDAS. La stima può essere fatta in più passi per matrici di dimensione elevata.*

Key words: Realized covariance, Realized DCC, Component models, MIDAS

Luc Bauwens

Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE), 34, Voie du Roman Pays B-1348 Louvain-la-Neuve (Belgium), e-mail: luc.bauwens@uclouvain.be

Manuela Braione

Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE), 34, Voie du Roman Pays B-1348 Louvain-la-Neuve (Belgium), e-mail: manuela.braione@uclouvain.be

Giuseppe Storti

Università di Salerno, Dipartimento di Scienze Economiche e Statistiche, Via Giovanni Paolo II, 132, 84084 Fisciano (Italy), e-mail: storti@unisa.it

1 Introduction

In recent years financial econometrics literature has been paying growing attention to the modelling of time series of realized covariance matrices. A review of these contributions can be found in Bauwens et al. (2012). Standard financial applications require the prediction of vast-dimensional covariance matrices leading to the rise of two relevant problems related to the number of parameters to be estimated and to the positive definiteness of the estimated conditional covariance matrices. First, at the model building stage, to limit the inflation of the number of estimated parameters, severe restrictions on the model dynamics must be imposed. Also, in order to further limit the number of parameters to be estimated, *targeting* procedures should be applied. In addition the chosen parameterization must guarantee the positive definiteness of the estimated covariance matrices.

The models that have been so far proposed imply that realized covariances are reverting to a constant long run level which is, in many settings, a questionable assumption. Under this respect, several papers provide empirical evidence in this direction proposing structural models that decompose volatility in two components related to long and short run movements. In the classical time series literature on GARCH models, examples of such models for univariate volatilities include the Spline-GARCH model by Engle and Rangel (2008) and, more recently, the GARCH-MIDAS model by Engle et al. (2013). Both these models identify a long term volatility component separated from a short term one and, in both cases, the two components are combined according to a multiplicative scheme. The main difference between these two approaches lies in the way the long term component is specified. In the first case it is modelled as a spline while in the latter model a MIDAS (see e.g. Engle et al., 2013) structure is assumed. In the multivariate framework, component volatility models have been proposed by Hafner and Linton (2010) and Colacito et al. (2011). The first model assumes a multiplicative structure, similar to that mentioned for the univariate case, where the long term component is estimated by a matrix-variate kernel smoother. Differently, the Component DCC model by Colacito et al. (2011) assumes a GARCH-MIDAS structure for the univariate volatilities while, for conditional correlations, the long term component is introduced as a time varying intercept, following a MIDAS model, in the dynamic updating equation for the rescaled conditional correlation.

In this paper our contribution is to propose some novel model specifications that allow to decompose the dynamics of realized covariance matrices into different components corresponding to different time frequencies. Section 2 illustrates the proposed models and the associated inference procedures while Section 3 reports the results of an application to a simulated dataset. Section 4 concludes.

2 Dynamic component models for realized covariance matrices

2.1 The general framework

Let C_t be a sequence of a positive definite and symmetric (PDS) realized covariance matrices of order n , for $t = 1, \dots, T$. We assume that conditional on past information I_{t-1} consisting of C_τ for $\tau \leq t-1$, and for all t , C_t follows a n -dimensional central Wishart distribution and denote this assumption by

$$C_t | I_{t-1} \sim W_n(\nu, S_t / \nu), \quad (1)$$

where $\nu (> n-1)$ is the degrees of freedom parameter and S_t / ν is the PDS scale matrix of order n . Equation (1) defines a generic conditional autoregressive Wishart (CAW) model, as proposed by Golosnoy et al. (2012). From the properties of the Wishart distribution it follows that

$$E(C_t | I_{t-1}) := E_{t-1}(C_t) = S_t, \quad (2)$$

so that the i, j -th element of S_t is defined as the conditional covariance between returns on assets i and j , $cov(r_{i,t}, r_{j,t} | I_{t-1})$, for $i, j = 1, \dots, n$, $r_{i,t}$ denoting the logarithmic return on asset i between the ends of periods $t-1$ and t .

We can incorporate a long term component in the specification of the expected realized covariance matrix by further decomposing S_t as follows:

$$S_t = L_t S_t^* L_t' \quad (3)$$

$$S_t^* = D_t^* R_t^* D_t^*, \quad (4)$$

where $D_t^* = \{\text{diag}(S_t^*)\}^{1/2}$. The matrix S_t^* is the short term component of the conditional covariance matrix, and R_t^* is the corresponding short term correlation matrix, whereas L_t is a matrix square root of the long term component of the covariance matrix that will be denoted by M_t .

2.2 A multiplicative model with a nonparametric long term component

The long run component M_t can be specified in two different ways. First, inspired by the component approach introduced in multivariate GARCH models by Hafner and Linton (2010) for conditional covariances and by Bauwens et al. (2013) for conditional correlations, the long run component matrix is assumed to be a function of the rescaled time index:

$$M_t = L_t L_t' = M(t/T). \quad (5)$$

One can also assume that the matrix M_t is a function of an observable variable instead of rescaled time. Such a variable should be included in the information set I_{t-1} and be lagged, so that we should write then $M_t = M_t(x_{t-1})$. A market volatility index (such as the VIX) could be considered as a relevant variable for x_t .

For identification, we assume that $E(S_t^*) = I_n$. In order to introduce this constraint into the model, we need to enforce firstly that the dynamic equations for the diagonal elements of $S_{ii,t}^*$ (the short term conditional variances) are specified in such a way that these conditional variances are reverting to a unit value. For example, such an equation, is

$$S_{ii,t}^* = (1 - \gamma_i - \delta_i) + \gamma_i C_{ii,t-1}^* + \delta_i S_{ii,t-1}^*, \quad (6)$$

where $C_{ii,t-1}^*$ is the i -th diagonal element of C_t^* , defined in (9). Secondly, we must impose that the dynamic equation for R_t^* is specified in such a way that R_t^* is reverting to an identity matrix I_n

$$R_t^* = (1 - \alpha - \beta) + \alpha P_{t-1}^* + \beta R_{t-1}^*, \quad (7)$$

where

$$P_t^* = \{\text{diag}(C_t^*)\}^{-1/2} C_t^* \{\text{diag}(C_t^*)\}^{-1/2} \quad (8)$$

and

$$C_t^* = L_t^{-1} C_t (L_t^{-1})'. \quad (9)$$

The matrix C_t^* is the realized covariance matrix purged of its long term component, and the matrix P_t^* is the corresponding correlation matrix.

For large dimensions, the estimation can be performed more easily in three steps. In the first step, due to the nonparametric specification of M_t , this set of matrices can be consistently estimated following the approach described in Hafner and Linton (2010) and Bauwens et al. (2013). Let $\hat{M}_t$ denote such a consistent estimator of M_t , and $\hat{L}_t$ such that $\hat{L}_t \hat{L}_t' = \hat{M}_t$. An example is the Nadaraya-Watson estimator of M_t given by

$$\hat{M}_t(\tau) = \frac{\sum_{t=1}^T K_h(\frac{t}{T} - \tau) C_t}{\sum_{t=1}^T K_h(\frac{t}{T} - \tau)}, \quad (10)$$

where $\tau \in [0, 1]$, $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ being a kernel function and h the bandwidth parameter. This estimator uses the same bandwidth for all elements of the covariance matrix. Conditional variance and correlation parameters are then estimated in the second and third step, respectively, by Quasi Maximum Likelihood.

2.3 A multiplicative model with a MIDAS long term component

An alternative way of modelling the long run component of the conditional variances and covariances or correlations is via the mixed data sampling (MIDAS) approach. In particular, the long run time varying level matrix $M_t = L_t L_t'$ of (3) is assumed to have the following MIDAS parameterization proposed by Golosnoy et

al. (2012):

$$M_t = \Lambda + \theta \sum_{l=1}^L \phi_l(\omega) \bar{C}_{t,l}^{(m)}, \quad (11)$$

where $\bar{C}_{t,l}^{(m)} = \sum_{\tau=t-l}^{t-1} C_{\tau}^{(m)}$. The latter matrix can also be replaced by C_{t-l} (using a larger value of the lag order L). In (11) θ is a positive scalar parameter and $\phi_l(\omega)$ is an a priori specified weight function (e.g. a beta function), with weights adding up to one and depending on the parameter vector $\omega = (\omega_1 \ \omega_2)$ for beta weights². Λ is a symmetric and positive-definite matrix of constant parameters that must be estimated jointly with θ and ω . The model for R_t^* and the short run conditional variances can be specified in several ways, for example like in (6) and (7). The model's parameters can be estimated by maximizing a likelihood function based on the Wishart assumption. Indeed it can be shown that if the conditional expectation of C_t is correctly specified, at the true value of the parameters the score of the log-likelihood function is a martingale difference sequence (MDS), so that this function has a quasi-likelihood interpretation and the estimator is consistent independently from the Wishart assumption (assuming the usual other conditions).

One difficulty of the specification of M_t in (11) is the proliferation of parameters in large dimensions, due to the matrix Λ . To overcome this problem a targeting procedure could be applied. If θ is equal to zero, the long term component is constant. In this case targeting can be performed by an approach similar to that pursued by Bauwens et al. (2012). Differently, when $\theta \neq 0$, an iterative procedure is required. If consistent estimators $\hat{\theta}$, $\hat{\omega}$, and $\hat{M}_t$ were available, then

$$\hat{\Lambda} = T^{-1} \sum_{t=1}^T \left(\hat{M}_t - \hat{\theta} \sum_{l=1}^L \phi_l(\hat{\omega}) \bar{C}_{t,l}^{(m)} \right) \quad (12)$$

could be used as a targeting estimator of Λ . This issue is currently left for future research.

3 Empirical results

In this section, in order to assess the effectiveness of the proposed approach for capturing slowly varying long-run volatility components, we present the results of a simulation exercise. In particular we simulate a single series of $T = 5000$ observation from a bivariate multiplicative component model where the long run component is given by a deterministic function specified, following Engle (2002), as a constant,

¹ See equation (28) in Golosnoy et al. (2012). For example, $\bar{C}_{t,1}^{(22)}$ is the realized covariance matrix computed over m days from day $t - 22$ to day $t - 1$, $\bar{C}_{t,2}^{(22)}$ from day $t - 44$ to $t - 23$, etc.

² The beta weight function, normalized so that the sum of the weights is equal to one, is equal to $\phi_l(\omega) = [(l/L)^{\omega_1-1} (1-l/L)^{\omega_2-1}] / [\sum_{j=1}^L (j/L)^{\omega_1-1} (1-j/L)^{\omega_2-1}]$, where ω_1 and ω_2 are both positive parameters and determine the shape of the weight function.

a step, a sine and a fast sine function. We have then fitted both the non-parametric and the MIDAS component models to the simulated data. Fig. 1 reports the simulated conditional variance for one of the process components together with the long run components estimated by the MIDAS and non-parametric approach. The results show that in both cases the estimated long-run components are able to closely follow the simulated trend.

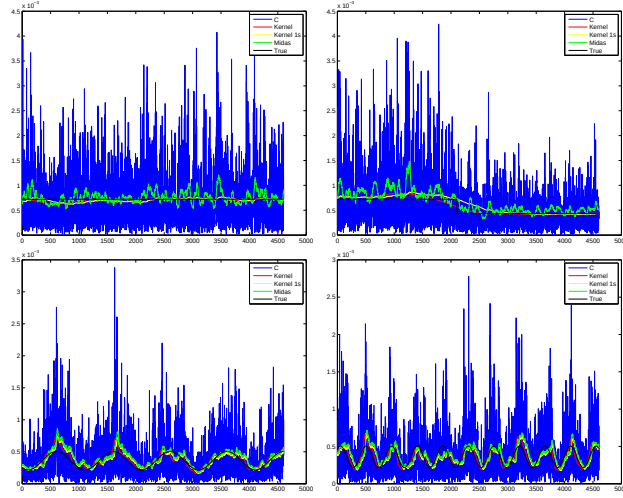


Fig. 1 Simulated and estimated volatilities from a multiplicative model ($T=5000$, $n=2$). K1S=kernel one sided, Kernel=kernel two-sided, C=simulated data, MIDAS=MIDAS estimate, True=simulated trend.

4 Concluding remarks

In this paper some structural models for realized covariance matrices have been proposed. The results of an empirical example show that the model are able to satisfactorily reproduce a wide range of long-run behaviours. Our projects for future research include the derivation of a targeting procedure for the estimation of large dimensional models. Also an additive component models is currently under study.

Acknowledgements Giuseppe Storti gratefully acknowledges financial support from MIUR within the PRIN project 2010- 2011 (prot. 2010J3LZEN): *Forecasting economic and financial time series: understanding the complexity and modelling structural change*.

References

1. Bauwens, L., Hafner, C., Pierret, D. (2013) Multivariate volatility modeling of electricity futures, *Journal of Applied Econometrics*, 28, 5, 743–761.
2. Bauwens, L., Storti, G., Violante, F. (2013) Dynamic Conditional Correlation Models for Realized Covariance Matrices, *CORE DP 2012/60*.
3. Colacito, R., Engle, R.F., Ghysels, E. (2011) A component model for dynamic correlations, *Journal of Econometrics*, 164, 1, 45–59.
4. Engle, R. F. (2002) *Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models*. *Journal of Business & Economic Statistics*, 20, 3, 339–350.
5. Engle, R.F., Ghysels, E., Sohn, B. (2013) Stock market volatility and macroeconomic fundamentals, *Review of Economic and Statistics*, 95, 3, 776–797.
6. Engle, R., Rangel, J.G. (2008) The spline-GARCH model for low-frequency volatility and its global macroeconomic causes, *Review of Financial Studies*, 21, 1187–1222.
7. Golosnoy, V., Gribisch, B., Liesenfeld, R. (2011) The Conditional Autoregressive Wishart Model for Multivariate Stock Market Volatility, *Journal of Econometrics*, 167, 211–223.
8. Hafner, C., Linton, O.B. (2010) Efficient estimation of a multivariate multiplicative volatility model, *Journal of Econometrics*, 159, 1, 55–73.