2012/28

Computationally efficient inference procedures
for vast dimensional realized covariance models

Luc Bauwens and Giuseppe Storti

CORE

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium

<http://www.uclouvain.be/core>

CORE DISCUSSION PAPER
2012/28

**Computationally efficient inference procedures
for vast dimensional realized covariance models**

Luc BAUWENS¹ and Giuseppe STORTI²

June 2012

Abstract

This paper illustrates some computationally efficient estimation procedures for the estimation of vast dimensional realized covariance models. In particular, we derive a Composite Maximum Likelihood (CML) estimator for the parameters of a Conditionally Autoregressive Wishart (CAW) model incorporating scalar system matrices and covariance targeting. The finite sample statistical properties of this estimator are investigated by means of a Monte Carlo simulation study in which the data generating process is assumed to be given by a scalar CAW model. The performance of the CML estimator is satisfactory in all the settings considered although a relevant finding of our study is that the efficiency of the CML estimator is critically dependent on the implementation settings chosen by modeller and, more specifically, on the dimension of the marginal log-likelihoods used to build the composite likelihood functions.

Keywords: realized covariance, CAW model, BEKK model, composite likelihood, covariance targeting, Wishart distribution.

JEL classification: C32, C58

¹ Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: luc.bauwens@uclouvain.be. This author is also member of ECORE, the association between CORE and ECARES.

² Department of Economics and Statistics, Universiti Salerno, Fisciano, Italy. E-mail: storti@unisa.it

Research supported by the contract "Projet d'Actions de Recherche Concertées" 07/12-002 of the "Communauté française de Belgique", granted by the "Académie universitaire Louvain". The scientific responsibility is assumed by the authors.

1 Introduction

Many financial applications, such as portfolio optimization or risk management, require to work with large dimensional portfolios involving a number of assets in the order of one hundred or even more. In order to obtain parsimonious multivariate volatility models, whose estimation is feasible in large dimensions, it is necessary to impose drastic homogeneity restrictions on the dynamic laws determining the evolution of conditional variances and covariances. However, even for parsimonious models, for very large dimensions the computation of (quasi) maximum likelihood estimates can be computationally challenging and troublesome. Also it is important to note that, for moderately large values of the model's dimension n (as a rule of the thumb, say for $50 \leq n \leq 100$), even if direct maximization of the likelihood or quasi likelihood function is feasible, the computational time required can be so high to prevent the use of resampling and simulation techniques and, in general, any application that implies the need to iteratively estimate the model for a large number of times. The application of bootstrap to realized covariance models has been recently investigated by [5] while the application of simulation based inference procedures is usually required for long-term prediction of risk measures such as Value at Risk and Expected Shortfall.

This has stimulated the research on the development of alternative algorithms for the generation of computationally efficient consistent estimators of the parameters of vast dimensional multivariate volatility models. These algorithms have been first developed for the estimation of Multivariate GARCH (MGARCH) models, see [2] for a review, but they can be modified or adapted, as shown in [3], for their application to models for realized covariance matrices.

An obvious approach to deal with inference in vast dimensional systems is to split the multivariate estimation problem into a set of simpler low-dimensional problems. This is the spirit of the McGyver method proposed by [7]. The algorithm is illustrated for the case of Dynamic Conditional Correlation (DCC) models with correlation targeting ([6]) although it can be readily applied to any *scalar* MGARCH model such as a scalar BEKK model ([8])¹. The basic idea is that, if the process dynamics are characterized by scalar parameters, these can be consistently estimated by fitting the model even to an arbitrarily chosen bivariate subsystem. The estimation can then be repeated over all the possible $n(n-1)/2$ different bivariate subsystems or a subset of them. The final estimate of model parameters is obtained by calculating the mean or median of the empirical distribution of the estimated coefficients. This estimate is expected to be more efficient than an estimate obtained by fitting the model to a single bivariate subsystem since it is implicitly pooling information from all the assets in the dataset. Evidence in this direction is provided by [7]. The procedure is in its nature heuristic but it is straightforward to show that it automatically returns a consistent estimator. However, [7] does not provide any analytical results on the asymptotic properties of the estimator, including its asymptotic distribution

¹We indicate as *scalar* any model in which all the conditional covariances or correlations share the same parameters

and efficiency. On the other hand, the finite sample statistical properties of the McGyver estimator are investigated by Monte Carlo simulation. One point which is left unexplored by [7] is related to the sensitivity of the properties of the estimation procedure to the size of the subsystems involved in its implementation.

An alternative approach to the estimation of vast dimensional conditional heteroskedastic models is based on Composite Likelihood theory (see [14] for a recent review). This approach replaces the full log-likelihood with an approximation based on the sum of low-dimensional log-likelihoods of a given dimension $m \ll n$. As for the McGyver method, the researcher can consider the full set of m -dimensional log-likelihoods or a subset of these. In [9] the CML approach is applied to the estimation of scalar DCC models with correlation targeting. From a computational point of view, the calculation of CML estimators is much faster than that of standard ML estimators. [9] also analyze different variants of the Composite Maximum Likelihood (CML) estimator and assess their finite sample properties by means of an extensive Monte Carlo study. Their findings show that, in large systems, the CML estimator can be much more efficient than the Maximum Likelihood (ML) estimator with correlation targeting. In particular, the simulation results reveal that this estimator tends to be affected by a systematic bias component whose size is increasing with n . In the paper this difference is ascribed to the high number of nuisance parameters involved in the optimization. Also, they show that the CML estimator favorably compares with the McGyver method being, by far, more efficient than the latter. In the light of these results, in this paper, our attention will be focused on CML estimation.

Both the CML and the McGyver method reformulate the estimation problem in terms of simpler lower dimensional problems but while, in the McGyver method, the final estimate is obtained as a function of the results of several low-dimensional optimizations, in CML estimation the optimization is performed just once leading to a substantial reduction of the computing time. Except for the recent contribution by [3], to our knowledge, so far this inference procedure has not been applied to the estimation of models for realized covariances. In this paper we will illustrate an approach to the estimation of large dimensional realized covariance models based on the maximization of a CL function derived under the assumption of Wishart marginal log-likelihoods. By means of a Monte Carlo simulation study we will i) evaluate the efficiency of the estimator in finite samples ii) investigate the sensitivity of the estimator's performance, bias and efficiency, to the size of the marginal log-likelihoods used to build the CL function and to the number of low-dimensional subsystems used in the computation.

Among the several different models for realized covariance matrices that have been proposed in the past years, we will focus on the class of Conditional Autoregressive Wishart (CAW) models recently introduced by [10]. These models are based on the assumption that the conditional distribution of the realized covariance matrix is a Wishart distribution with time varying conditional expectation proportional to the scale matrix of the Wishart. In the basic version of the model the time varying conditional expectation of the realized covariance matrix is assumed to follow a BEKK ([8]) type specification. Unless further restrictions on the parameter space are imposed, this assumption still leaves the number of parameters linear in n^2 where n is the number of assets considered. In their paper, [10]

present an application to a dataset including 5 stocks where the estimated model includes 116 parameters. It is easy to understand how fitting an unconstrained CAW model to a dataset whose dimension is even moderately large is not feasible. Hence, in this paper, we will consider a restricted version of the CAW model in which the parameter matrices of the dynamic equation for the conditional expectation of the realized covariance matrix are assumed to be scalar. A covariance targeting approach will be also used in order to avoid direct estimation of the matrix of intercepts of the BEKK recursion.

The structure of the paper is as follows. In section 2 we will illustrate the CAW model and discuss its statistical properties. Maximum likelihood inference for the CAW model will be discussed in section 3 while, in section 4, we will present the CML estimator for the parameters of a restricted CAW model. Section 5 will report the results of a Monte Carlo simulation study and conclude.

2 The Conditional Autoregressive Wishart (CAW) model

The CAW model, recently proposed by [10], is based on the assumption that the conditional distribution of realized covariance matrices, given past information, is Wishart. In the literature on realized covariance models the Wishart assumption is not new and has already been used in a number of papers, starting from [11] who were probably the first to use the idea of a Wishart process for realized covariance matrices with the WAR(p) (Wishart autoregressive) process, where p is a lag order parameter. They assume that the conditional distribution of realized covariance matrices follows a non-central Wishart distribution where the matrix of non-centrality parameters is modelled as a function of past lagged realized covariance matrices. Practical application of this model has been limited by the fact that it is too heavily parameterized since it uses a number of parameters equal to $3n^2/2 + n/2 + 1$ (for the one lag case). In order to obtain a more parsimonious model structure, [4] propose a block structured version of the WAR model which significantly reduces the number of parameters but still keeps it of the order of n^2 . A different approach is taken by [12] and [13] who propose to jointly model the returns vector and its realized covariance matrix. In both papers, just as in the CAW model, the realized covariance part of the model specifies the conditional distribution of the realized covariance as a Wishart, whose expected value is proportional to the scale matrix of the Wishart. In [12], that scale matrix is modelled as a function of a few lags of itself while, in [13], it is assumed to follow a BEKK process. The CAW model shares with [13] the assumption of conditional Wishart distribution and BEKK type dynamics for the realized covariance models.

More specifically, let C_t , for $t = 1, \dots, T$, be a time series of positive definite symmetric (PDS) realized covariance matrices of dimension n . It is assumed that the conditional distribution of C_t , given past information on the history of the process I_{t-1} , consisting of C_τ for $\tau \leq t-1$, and $\forall t$, is given by a n -dimensional central Wishart distribution

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

where ν ($> n - 1$) is the degrees of freedom parameter and S_t/ν is a PDS scale matrix of order n . From the properties of the Wishart distribution (see [1], among the others) it follows that

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

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 . Equation (1) defines a generic conditional autoregressive Wishart (CAW) model as proposed by [10]. The specification of the dynamic updating equation for S_t can be chosen within a wide range of options. Namely [10] use a BEKK-type formulation mutuuted from the MGARCH literature. For a model of order (p, q) , this corresponds to the following dynamic equation

$$S_t = GG' + \sum_{i=1}^p A_i C_{t-i} A_i' + \sum_{j=1}^q B_j S_{t-j} B_j' \quad (3)$$

where A_i and B_j are square matrices of order n and G is a lower triangular matrix such that GG' is PDS. This choice also ensures that S_t is PDS for all t if S_0 is itself PDS. In order to guarantee model identifiability, it is sufficient to assume that the main diagonal elements of G and the first diagonal element for each of the matrices A_i, B_j are positive. There are two main differences between the WAR model of [11] and the CAW model: i) it is assumed that the conditional distribution of C_t is a central Wishart distribution rather than a non-central one ii) the CAW model analyzes the dynamic evolution of the scale matrix S_t while the WAR model is focusing on the matrix of non-centrality parameters.

In order to investigate the statistical properties of the CAW(p,q) model it is useful to consider two alternative observationally equivalent representations of the model. First, The CAW(p,q) model in (1-3) can be represented as a state-space model with observation equation given by

$$C_t = \frac{1}{\nu} S_t^{1/2} U_t (S_t^{1/2})' \quad U_t \sim W_n(\nu, I_n); \quad (4)$$

where $S_t^{1/2}$ denotes the lower triangular matrix obtained from the Cholesky factorization of S_t such that $S_t^{1/2} (S_t^{1/2})' = S_t$ and U_t is a measurement error distributed as a standardized Wishart distribution with identity scale matrix and degrees of freedom equal to ν . S_t acts as a matrix-variate state variable with dynamic transition equation given by (3). This representation allows to interpret S_t as the “true” latent integrated covariance for a wide class of multivariate continuous time stochastic volatility models, C_t as a consistent estimator of S_t and U_t as the associated matrix of estimation errors. Second, a CAW(p,q) process admits the following VARMA representation

$$c_t = g + \sum_{i=1}^{\max(p,q)} (\mathcal{A}_i + \mathcal{B}_i) c_{t-i} + \sum_{j=1}^q \mathcal{B}_j v_{t-j} + v_t \quad (5)$$

where $g = \text{vech}(GG')$ and $c_t = \text{vech}(C_t)$; v_t is a martingale difference error term such that $E(v_t) = 0$ and $E(v_t v_s') = 0$, $\forall t \neq s$; $\mathcal{A}_i = L_{n*} (A_i \otimes A_i) D_{n*}$ and $\mathcal{B}_j = L_{n*} (B_j \otimes B_j) D_{n*}$, with

$n^* = n(n+1)/2$ ($\mathcal{A}_i = 0$ for $i > p$, $\mathcal{B}_j = 0$ for $j > q$). L_{n^*} and D_{n^*} are duplication and elimination matrices such that $\text{vec}(X) = D_{n^*} \text{vech}(X)$ and $\text{vech}(X) = L_{n^*} \text{vec}(X)$. This representation allows to derive conditions for the existence of the unconditional mean of the CAW(p,q) process. In particular, it can be shown that $\bar{c} = E(c_t)$ will be finite if and only if all the eigenvalues of the matrix $\Psi = \mathcal{A}_i + \mathcal{B}_j$ are less than 1 in modulus. In this case we will have

$$\bar{c} = E(c_t) = (I_{n^*} - \sum_{i=1}^{\max(p,q)} (\mathcal{A}_i + \mathcal{B}_i))^{-1} g \quad (6)$$

For large n , the model formulation in (3) renders the estimation unfeasible due to the high number of parameters. As we are interested in applying the model to situations in which n is large (say $n \approx 50$) we consider a modified version of equation (3) in which the parameter matrices A_i and B_j are replaced by scalars and covariance targeting is used to avoid simultaneous estimation of the matrix G

$$S_t = (1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j) \bar{C} + \sum_{i=1}^p \alpha_i C_{t-i} + \sum_{j=1}^q \beta_j S_{t-j} \quad (7)$$

with $\bar{C} = E(C_t)$, $\alpha_i = a_i^2$ and $\beta_j = b_j^2$. The particular implementation of covariance targeting used in this formulation is justified by the fact that, applying equation (6) to model (7) returns

$$\bar{C} = \frac{GG'}{(1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j)}.$$

In practice, covariance targeting is implemented by consistently pre-estimating the unconditional expectation of C_t by the sample average $\hat{\bar{C}} = T^{-1} \sum_{t=1}^T C_t$ and substituting this estimator for the corresponding population moment in (7). The scalar model implies that the conditional variances and covariances all follow the same dynamic pattern which is indeed a restrictive assumption. However, this compromise is necessary in order to obtain a parsimonious model whose estimation is tractable in high dimensions. A generalization of this framework is discussed by [3].

3 Quasi Maximum Likelihood estimation of CAW models

In their paper [10] perform the estimation of model parameters by the method of maximum likelihood. The conditional density of C_t given past information is

$$f(C_t | I_{t-1}) = \frac{|S_t/\nu|^{-\nu/2} |C_t|^{(\nu-n-1)/2}}{2^{\nu n/2} \pi^{n(n-1)/4} \prod_{i=1}^n \Gamma\{(\nu+1-i)/2\}} \exp \left\{ -\frac{1}{2} \text{tr}(\nu S_t^{-1} C_t) \right\}.$$

It follows that, up to a constant, the log-likelihood contribution of observation t is given by

$$\ell_t(\theta) = \lambda(\nu) - \frac{\nu}{2} \log(|S_t|) - \frac{\nu}{2} \text{tr}(S_t^{-1} C_t) \quad (8)$$

where

$$\lambda(\nu) = \frac{\nu n}{2} \log(\nu) + \frac{\nu - n - 1}{2} \log(|C_t|) - \frac{\nu n}{2} \log(2) - \sum_{i=1}^n \log[\Gamma\{(\nu + 1 - i)/2\}].$$

The last two terms on the right hand side of (8) are proportional to the value of the Wishart shape parameter ν . It immediately follows that the first order conditions for the estimation of a_i and b_j ($i = 1, \dots, p; j = 1, \dots, q$), the parameters of the dynamic updating equation for C_t , do not depend on ν since the score for observation t with respect to $\theta_c = (a_1, \dots, a_p, b_1, \dots, b_q)'$ is given by

$$\frac{\partial \ell_t}{\partial \theta_c} = -\frac{\nu}{2} \left\{ \frac{\partial \log(|S_t|)}{\partial \theta_c} + \frac{\partial \text{tr}(S_t^{-1} C_t)}{\partial \theta_c} \right\}. \quad (9)$$

This implies that the value of ν is not affecting the estimation of θ_c that can be consistently estimated independently of the value of this parameter. We also note, similarly to [3], that, under the assumption that the dynamic model for the conditional expectation of C_t is correctly specified, the score in equation (9) is a martingale difference sequence even if the Wishart assumption is not satisfied. Analytically, the derivative vector in (9) is given by

$$\frac{\partial \ell_t}{\partial \theta_c} = -\frac{\nu}{2} \left\{ \text{tr} \left(S_t^{-1} \frac{\partial S_t}{\partial \theta_c} \right) + \text{tr} \left(C_t S_t^{-2} \frac{\partial S_t}{\partial \theta_c} \right) \right\}. \quad (10)$$

Taking expectations of both sides of (10), conditional on past information I_{t-1} , gives

$$\begin{aligned} E \left(\frac{\partial \ell_t}{\partial \theta_c} | I_{t-1} \right) &= \frac{\nu}{2} \left\{ \text{tr} \left(S_t^{-1} \frac{\partial S_t}{\partial \theta_c} \right) + \text{tr} \left(E(C_t | I_{t-1}) S_t^{-2} \frac{\partial S_t}{\partial \theta_c} \right) \right\} \\ &= \frac{\nu}{2} \left\{ \text{tr} \left(S_t^{-1} \frac{\partial S_t}{\partial \theta_c} \right) + \text{tr} \left(S_t^{-1} \frac{\partial S_t}{\partial \theta_c} \right) \right\} \end{aligned} \quad (11)$$

At the true parameter value $\theta_c = \theta_{c,0}$, we have that $E(C_t | I_{t-1}) = S_t$ by application of equation (2). By substituting this in (11) we obtain

$$E \left(\frac{\partial \ell_t}{\partial \theta_c} | I_{t-1}, \theta_{c,0} \right) = \frac{\nu}{2} \left\{ \text{tr} \left(S_t^{-1} \frac{\partial S_t}{\partial \theta_c} \right) + \text{tr} \left(S_t^{-1} \frac{\partial S_t}{\partial \theta_c} \right) \right\} = 0.$$

Under the usual regularity conditions, this result allows to interpret $\hat{\theta}_c$, the maximizer, with respect to θ_c , of the log-likelihood obtained as the sum of (8), as a Quasi Maximum Likelihood (QML) estimator. Hence, even if the Wishart assumption on the conditional distribution of C_t is not satisfied, the resulting estimator can still be proven to be consistent and asymptotically normal (see [3] for details).

4 Composite Likelihood estimation of CAW models

As already discussed in section 1, practical computation of the ML estimator for large dimensional models can be highly computational intensive and, for very large models, even unfeasible. The main reason for this is the necessity of iteratively inverting the high dimensional covariance matrix S_t in the log-likelihood function at each observation and each iteration of the optimization procedure. Given the values of the time series (T) and cross-sectional (n) dimensions typically considered in multivariate volatility modelling, this operation can be very time consuming. CML estimation offers a practical solution to this problem since the log-likelihood function is approximated by the sum of many low-dimensional marginal log-likelihood functions. In order to face a similar issue arising in the estimation of DCC models, [9] propose to apply the CML method under the assumption of conditionally normal returns. Differently, in the setting which is being here considered, we derive a CL function for the realized covariance matrix under the assumption of a conditionally Wishart distribution. The derivation of the CL function relies on two results. Before proceeding with their illustration we first need to define the following notations. For any square matrix M_t of order n , we denote by $M_{AA,t}$ a square matrix of order n_A extracted from M_t , which has its main diagonal elements on the main diagonal of M_t . Namely, if A stands for a subset of n_A different indices of $1, 2, \dots, n$, $M_{AA,t}$ is the matrix that consists of the intersection of the rows and columns of M_t corresponding to the selection of indices denoted by A . The two results can then be formulated as:

R1 : If $C_t \sim W_n(\nu, S_t/\nu)$, for any selection of n_A indices we have

$$C_{AA,t} \sim W_{n_A}(\nu, S_{AA,t}/\nu);$$

R2 : If $S_t = (1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j) \bar{C} + \sum_{i=1}^p \alpha_i C_{t-i} + \sum_{j=1}^q \beta_j S_{t-j}$,

$$S_{AA,t} = (1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j) \bar{C}_{AA} + \sum_{i=1}^p \alpha_i C_{AA,t-i} + \sum_{j=1}^q \beta_j S_{AA,t-j}.$$

Result 1 is a well known property of the Wishart distribution. By properties of the Wishart distribution, the marginal distribution of $C_{AA,t}$ is also Wishart, with the same degrees of freedom and with scale matrix obtained by deleting from S_t the same rows and columns as in C_t - see [1], Theorem 7.3.4. Applying this result with $n_A = 1$ corresponds to the result that the margin of a diagonal element of a Wishart matrix is a Gamma distribution. Result 2 is an obvious algebraic result.

The CML estimator of the parameters α and β is then defined as the maximizer of the sum of a number of Wishart marginal log-likelihoods for sub-matrices $C_{AA,t}$ corresponding to different choices of indices A . The simplest choice is to select all the log-likelihoods corresponding to sub-matrices of order 2, i.e. to all the $n(n-1)/2$ covariances or pairs of assets. Notice that with these bivariate Wishart log-likelihoods, only matrices of order 2 must be inverted, which can be efficiently programmed. We will denote any Wishart CL

functions based on marginal log-likelihoods of dimension 2 as CL_2 . Formally, this leads to the following expression

$$CL_2(\nu, \theta_c, \hat{C}) = \sum_{h=2}^n \sum_{k < h} \ell_{hk,t}(\nu, \theta_c, \hat{C}_{hk}) \quad (12)$$

with

$$\begin{aligned} \ell_{hk,t}(\cdot) = & \frac{\nu}{2} \log(\nu) + \frac{\nu-3}{2} \log(|C_t^{(hk)}|) - \frac{\nu}{2} \log(2) - \sum_{i=1}^2 \log[\Gamma\{(\nu+1-i)/2\}] \\ & - \frac{\nu}{2} \log(|S_t^{(hk)}|) - \frac{\nu}{2} \text{tr}\{(S_t^{(hk)})^{-1} C_t^{(hk)}\}, \end{aligned} \quad (13)$$

where for any matrix M_t , $M_t^{(hk)}$ is the matrix of order 2 extracted at the intersection of rows h and k of M_t . In principle, one can use less terms than the $n(n-1)/2$ terms in (12) without affecting the consistency of the estimator. This can be particularly useful in particular in cases in which n is large. At the price of a slightly more complicated notation the expression in (13) can be easily generalized to the case in which marginal log-likelihoods of sub-matrices of higher dimension ($m \geq 2$) are used. Under this regard it is important to note that the number of subsystems that can be created, differing for at least one asset, is rapidly increasing with m , making the estimation problem soon unfeasible. The problem is illustrated in Table 1 for different values of m and n . For $n=50$ assets we

$n \setminus m$	2	3	4	5
10	45	120	210	252
25	300	2300	12650	53130
50	1225	19600	230300	2118760
75	2775	67525	1215450	17259390
100	4950	161700	3921225	75287520

Table 1: Number of m - dimensional subsystems versus cross-sectional dimension n .

have 1225 bivariate subsystems but 19600 trivariate subsystems which are different for at least one asset. In the case of $n=100$ assets, these numbers increase up to 4950 for $m=2$ and to 161700 for $m=3$. These values further increase for $m > 3$. It follows that for $m > 2$ the Composite Likelihood function can be practically built only using a subsample of all the available marginal likelihoods. Reducing the number of subsystems upon which the composite likelihood is based is however expected to reduce efficiency as empirically shown by [9]. On the other hand increasing the value of m is expected to increase efficiency. For $m = n$, we recover the maximum likelihood estimator as a special case.

5 Monte Carlo simulation

In this section we present the results of a Monte Carlo simulation study aimed at evaluating and comparing the finite sample efficiencies of the CML estimator derived under different implementation settings. In particular we will focus on the analysis of CML estimators based on the use of bivariate (CL2) and trivariate marginals (CL3), respectively. The data generating process is assumed to be given by a CAW(1,1) process of dimension $n = 50$ with covariance targeting, as defined by equations (2) and (7) and replacing $\bar{C}$ by the sample average $\hat{\bar{C}}$, and parameters given by $\alpha_1 \equiv \alpha = 0.05$, $\beta_1 \equiv \beta = 0.90$ and $\nu = n$. We generate 500 time series of length $T = 2000$ and $T = 3500$ respectively. In both cases the first 500 data points are discarded in order to reduce the impact of initial conditions.

In the bivariate case we compare the estimator based on all the potential bivariate subsystems with three alternatives in which the CL function is built from a subset of the universe of available bivariate systems. First, we consider a subset of dimension $(n - 1)$ composed by the set of contiguous pairs $\{i, i + 1\}$, for $i = 1, \dots, n - 1$. The rationale behind this choice is that in this way CML estimates are based on the minimal set guaranteeing that all assets are equally represented in the estimation process. Second, we consider two randomly selected subsets of dimension $[M/3]$ and $[M/2]$ respectively, where $[.]$ denotes rounding to the closest integer and M is the overall number of available subsystems. In the trivariate case, considering the whole set of different trivariate systems is not computationally feasible. Hence, the maximum number of trivariate systems which has been considered has been set equal to $N_m = 5000$. Again, as in the bivariate case, we analyze the sensitivity of the efficiency of the estimator with respect to the number of subsystems used for estimation by considering alternative estimators based on the set of $(n - 2)$ contiguous triplets and on random sets of size $[N_m/3]$ and $[N_m/2]$, respectively. The results have been summarized in figures 1, for α , and 2, for β . Both *CL2* and *CL3* result to be approximately unbiased even for the shorter sample size. As far as efficiency of the estimators is concerned, in comparative terms, it is evident that CL3 is remarkably more efficient than CL2 while it appears that, within the range of values considered for the simulation, the number of lower dimensional subsystems used to compute the CL function is not dramatically affecting the efficiency of the CL estimators. In particular, the efficiency gap between CL estimators based on contiguous systems and more complex estimators considering all the systems (bivariate case) or a large random sample of these (trivariate case) appears not substantial. In absolute terms, the efficiency level of both estimators is reasonably high (Table 2). The variation coefficient of CL estimators for parameter α is slightly lower than 4%, for CL2, and approximately equal to 2.5% for CL3 while, for parameter β , the recorded values are approximately equal to 0.5% for CL2 and to 0.3% for CL3.

In this paper only empirical (simulation based) results have been presented. From a theoretical point of view, an issue which has not been addressed is the derivation of the asymptotic distribution of the CL estimators and an estimator of their asymptotic standard errors. Given the complexity of the CAW model, this is not an easy task and we plan to investigate this issue in our future research.

	B1	B2	B3	B4	T1	T2	T3	T4
$\hat{s}e(\hat{a})/\hat{\mu}(\hat{a})$	0.0392	0.0388	0.0388	0.0386	0.0257	0.0244	0.0246	0.0246
$\hat{s}e(\hat{b})/\hat{\mu}(\hat{b})$	0.0054	0.0052	0.0052	0.0052	0.0036	0.0033	0.0033	0.0033

Table 2: Simulated variation coefficient (simulated s.e. ($\hat{s}e$)/simulated mean ($\hat{\mu}$)) of parameter estimates for CML estimators based on bivariate (Bi) and trivariate (Ti) marginals, respectively.

References

- [1] Anderson, T.W.: An Introduction to Multivariate Statistical Analysis, Wiley, 2nd Edition (1984).
- [2] Bauwens, L., Laurent, S., Rombouts, J.V.K.: Multivariate GARCH Models: A Survey, Journal of Applied Econometrics, **21**, 79–109 (2006).
- [3] Bauwens, L., Storti, G., Violante, F.: Dynamic conditional correlation models for realized covariance matrices, to appear as CORE Discussion Paper (2012).
- [4] Bonato, M., Caporin, M., Rinaldo, A.: Forecasting realized (co)variances with a block structure Wishart autoregressive model, Working Papers 2009-03, Swiss National Bank (2009).
- [5] Dovonon, P., Goncalves, S., Meddahi, N.: Bootstrapping realized multivariate volatility measures, Available at SSRN: <http://dx.doi.org/10.2139/ssrn.1500028> (2009).
- [6] Engle, R.F.: Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models, Journal of Business and Economic Statistics, **20**, 339–350 (2002).
- [7] Engle, R. F.: High dimensional dynamic correlations, In J. L. Castle and N. Shephard (Eds.), The Methodology and Practice of Econometrics: Papers in Honour of David F Hendry. Oxford University Press (2009).
- [8] Engle, R., Kroner, F.K.: Multivariate Simultaneous Generalized ARCH, Econometric Theory, **11**, 122–150 (1995).
- [9] Engle, R.F., Shephard, N., Sheppard, K.: Fitting vast dimensional time-varying covariance models, NYU Working Paper No. FIN-08-009 (2008).
- [10] Golosnoy, V., Gribisch, B., Liesenfeld, R.: The Conditional Autoregressive Wishart Model for Multivariate Stock Market Volatility, Journal of Econometrics, **167**, 211–223 (2012).
- [11] Gouriéroux, C., Jasiak, J., Sufana, R.: The Wishart Autoregressive process of multivariate stochastic volatility, Journal of Econometrics, **150**, 167–181 (2009).

- [12] Jin, X., Maheu, J.M.: Modelling Realized Covariances and Returns, WP 11-08, The Rimini Center for Economic Analysis (2010).
- [13] Noureldin, D., Shephard, N., Sheppard, K.: Multivariate High-Frequency-Based Volatility (HEAVY) Models, Journal of Applied Econometrics, forthcoming (2012).
- [14] Varin, C., Reid, N., Firth, D: An Overview of Composite Likelihood Methods, Statistica Sinica, **21**, 5–42 (2011).

Figure 1: Simulation results for the estimation of the α parameter by CML with bivariate (B) and trivariate (T) marginals. Key to graph: 1= estimator based on contiguous sets; 2=estimator based on $m_1 = \lfloor N_m/3 \rfloor$ subsystems; 3=estimator based on $m_2 = \lfloor N_m/2 \rfloor$ subsystems; 4=estimator based on $m_3 = N_m$ subsystems where $N_m = \min(M, 5000)$ and $\lfloor \cdot \rfloor$ indicates rounding to the closest integer.

Figure 2: Simulation results for the estimation of the β parameter by CML with bivariate (B) and trivariate (T) marginals. Key to graph: 1= estimator based on contiguous sets; 2=estimator based on $m_1 = \lfloor N_m/3 \rfloor$ subsystems; 3=estimator based on $m_2 = \lfloor N_m/2 \rfloor$ subsystems; 4=estimator based on $m_3 = N_m$ subsystems where $N_m = \min(M, 5000)$ and $\lfloor \cdot \rfloor$ indicates rounding to the closest integer.

Recent titles

CORE Discussion Papers

- 2011/57. Knud J. MUNK. Optimal taxation in the presence of a congested public good and an application to transport policy.
- 2011/58. Luc BAUWENS, Christian HAFNER and Sébastien LAURENT. Volatility models.
- 2011/59. Pierre PESTIEAU and Grégory PONTIERE. Childbearing age, family allowances and social security.
- 2011/60. Julio DÁVILA. Optimal population and education.
- 2011/61. Luc BAUWENS and Dimitris KOROILIS. Bayesian methods.
- 2011/62. Florian MAYNERIS. A new perspective on the firm size-growth relationship: shape of profits, investment and heterogeneous credit constraints.
- 2011/63. Florian MAYNERIS and Sandra PONCET. Entry on difficult export markets by Chinese domestic firms: the role of foreign export spillovers.
- 2011/64. Florian MAYNERIS and Sandra PONCET. French firms at the conquest of Asian markets: the role of export spillovers.
- 2011/65. Jean J. GABSZEWICZ and Ornella TAROLA. Migration, wage differentials and fiscal competition.
- 2011/66. Robin BOADWAY and Pierre PESTIEAU. Indirect taxes for redistribution: Should necessity goods be favored?
- 2011/67. Hylke VANDENBUSSCHE, Francesco DI COMITE, Laura ROVEGNO and Christian VIEGELAHN. Moving up the quality ladder? EU-China trade dynamics in clothing.
- 2011/68. Mathieu LEFEBVRE, Pierre PESTIEAU and Grégory PONTIERE. Measuring poverty without the mortality paradox.
- 2011/69. Per J. AGRELL and Adel HATAMI-MARBINI. Frontier-based performance analysis models for supply chain management; state of the art and research directions.
- 2011/70. Olivier DEVOLDER. Stochastic first order methods in smooth convex optimization.
- 2011/71. Jens L. HOUGAARD, Juan D. MORENO-TERNERO and Lars P. ØSTERDAL. A unifying framework for the problem of adjudicating conflicting claims.
- 2011/72. Per J. AGRELL and Peter BOGETOFT. Smart-grid investments, regulation and organization.
- 2012/1. Per J. AGRELL and Axel GAUTIER. Rethinking regulatory capture.
- 2012/2. Yu. NESTEROV. Subgradient methods for huge-scale optimization problems.
- 2012/3. Jeroen K. ROMBOUTS, Lars STENTOFT and Francesco VIOLANTE. The value of multivariate model sophistication: An application to pricing Dow Jones Industrial Average options.
- 2012/4. Aitor CALO-BLANCO. Responsibility, freedom, and forgiveness in health care.
- 2012/5. Pierre PESTIEAU and Grégory PONTIERE. The public economics of increasing longevity.
- 2012/6. Thierry BRECHET and Guy MEUNIER. Are clean technology and environmental quality conflicting policy goals?
- 2012/7. Jens L. HOUGAARD, Juan D. MORENO-TERNERO and Lars P. ØSTERDAL. A new axiomatic approach to the evaluation of population health.
- 2012/8. Kirill BORISOV, Thierry BRECHET and Stéphane LAMBRECHT. Environmental maintenance in a dynamic model with heterogeneous agents.
- 2012/9. Ken-Ichi SHIMOMURA and Jacques-François THISSE. Competition among the big and the small.
- 2012/10. Pierre PESTIEAU and Grégory PONTIERE. Optimal lifecycle fertility in a Barro-Becker economy.
- 2012/11. Catherine KRIER, Michel MOUCHART and Abderrahim OULHAJ. Neural modelling of ranking data with an application to stated preference data.
- 2012/12. Matthew O. JACKSON and Dunia LOPEZ-PINTADO. Diffusion and contagion in networks with heterogeneous agents and homophily.
- 2012/13. Claude D'ASPREMONT, Rodolphe DOS SANTOS FERREIRA and Jacques THEPOT. Hawks and doves in segmented markets: A formal approach to competitive aggressiveness.

Recent titles

CORE Discussion Papers - continued

- 2012/14. Claude D'ASPREMONT and Rodolphe DOS SANTOS FERREIRA. Household behavior and individual autonomy: An extended Lindahl mechanism.
- 2012/15. Dirk VAN DE GAER, Joost VANDENBOSSCHE and José Luis FIGUEROA. Children's health opportunities and project evaluation: Mexico's *Oportunidades* program.
- 2012/16. Giacomo VALLETTA. Health, fairness and taxation.
- 2012/17. Chiara CANTA and Pierre PESTIEAU. Long term care insurance and family norms.
- 2012/18. David DE LA CROIX and Fabio MARIANI. From polygyny to serial monogamy: a unified theory of marriage institutions.
- 2012/19. Carl GAIGNE, Stéphane RIOU and Jacques-François THISSE. Are compact cities environmentally friendly?
- 2012/20. Jean-François CARPANTIER and Besik SAMKHARADZE. The asymmetric commodity inventory effect on the optimal hedge ratio.
- 2012/21. Concetta MENDOLICCHIO, Dimitri PAOLINI and Tito PIETRA. Asymmetric information and overeducation.
- 2012/22. Tom TRUYTS. Stochastic signaling: Information substitutes and complements.
- 2012/23. Pierre DEHEZ and Samuel FEREY. How to share joint liability: A cooperative game approach.
- 2012/24. Pilar GARCIA-GOMEZ, Erik SCHOKKAERT, Tom VAN OURTI and Teresa BAGO D'UVA. Inequity in the face of death.
- 2012/25. Christian HAEDO and Michel MOUCHART. A stochastic independence approach for different measures of concentration and specialization.
- 2012/26. Xavier RAMOS and Dirk VAN DE GAER. Empirical approaches to inequality of opportunity: principles, measures, and evidence.
- 2012/27. Jacques H. DRÈZE and Erik SCHOKKAERT. Arrow's theorem of the deductible : moral hazard and stop-loss in health insurance.
- 2012/28. Luc BAUWENS and Giuseppe STORTI. Computationally efficient inference procedures for vast dimensional realized covariance models.

Books

- G. DURANTON, Ph. MARTIN, Th. MAYER and F. MAYNERIS (eds) (2010), *The economics of clusters – Lessons from the French experience*. Oxford University Press.
- J. HINDRIKS and I. VAN DE CLOOT (eds) (2011), *Notre pension en héritage*. Itinera Institute.
- M. FLEURBAEY and F. MANIQUET (eds) (2011), *A theory of fairness and social welfare*. Cambridge University Press.
- V. GINSBURGH and S. WEBER (eds) (2011), *How many languages make sense? The economics of linguistic diversity*. Princeton University Press.
- I. THOMAS, D. VANNESTE and X. QUERRIAU (eds) (2011), *Atlas de Belgique – Tome 4 Habitat*. Academia Press.
- W. GAERTNER and E. SCHOKKAERT (eds) (2012), *Empirical social choice*. Cambridge University Press.
- L. BAUWENS, Ch. HAFNER and S. LAURENT (eds) (2012), *Handbook of volatility models and their applications*. Wiley.
- J-C. PRAGER and J. THISSE (eds) (2012), *Economic geography and the unequal development of regions*. Routledge.

CORE Lecture Series

- R. AMIR (2002), Supermodularity and complementarity in economics.
- R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.
- A. SHAPIRO (2010), Stochastic programming: modeling and theory.