# Chain Graph for VAR and MARCH parameters reduction: EU Index returns case

# Grafi a Catena per la riduzione dei parametri VAR e MARCH: il caso degli indici EU

Andrea Pierini

**Abstract** In this paper a use of Chain Graph models is applied to a multivariate time series of EU index returns, in order to construct graphs with minimum BIC among a particular class of graphs called decomposable, which have the desirable property of a closed form estimation .
Then based on the previous identified relationships among present and past values of the multivariate time series, thanks to the chain graph modelling, a $VAR(1) - MArch(1)$ model is constructed by restricting to zero the parameters which are not indicated by the graphs. In this way a great reduction of parameters is put in place, using the opportune multidimensional modelling only if it is necessary.

In questo articolo i modelli grafici a catena sono usati al fine di costruire grafi di minimo BIC scelti tra i grafi decomponibili per i quali esiste una forma esplicita per la stima. Essi sono applicati ai ritorni multivariati degli indici EU. Sulla base delle relazioni identificate in precedenza tra il passato e il presente delle serie grazie i modelli grafici ai catena, si construisce un modello $VAR(1) - MArch(1)$ imponendo che i parametri non indicati dai grafi siano nulli. In tal modo si ottiene una notevole riduzione dei parametri, usando solo quelli necessari.

**Key words:** Chain Graph Model, VAR, MARCH, EU returns.

## 1 Introduction

The multiple financial data, as the stock Index returns, are treated in literature with heteroskedastic models (*MARCH*) , [1], and vector autoregressive models (*VAR*) which rests on the predictability of the average return on stocks, [2]. The

Andrea Pierini
University of Roma Tre, Department of Economics
Via Silvio D'Amico, 00145 Roma, e-mail: andrea.pierini@uniroma3.it

$VAR-MARCH$ type models suffer of computational burden as the number of parameters to be estimated are large with respect to univariate models. This paper suggests the use of Chain Graphs to identify only the significant parameters to be estimated for the $VAR-MARCH$ models whenever a subset of the variables are considered connected. The Chain Graphs are able to exploit the time ordering of the variables, that is some variables are prior in time to others. So the variables are divided into blocks and the modelling respects this ordering between this blocks without assuming any ordering within the blocks (time contemporaneity). Firstly a chain graph is built based on the returns and its immediate past.In case of connected variables, *Restricted* $VAR(1)$ models are estimated taking into account only the connections identified before. As the residuals exhibit heteroskedasticity, $MARCH(1)$ models are then applied to them in order to take this particular aspect into due consideration. A second chain graph is built based on the residuals and its immediate past.In case of connected residuals, *Restricted* $MARCH(1)$ models are estimated using only the connections previously identified. For the subset of variables which are not connected at all, univariate $AR(1)-Arch(1)$ are estimated. In this way a great reduction of calculation is obtained and a more efficient estimation is reached with respect to multidimensional full models.The orders $p,P$ of $VAR(p),Arch(P)$ can be augmented leaving the approach unchanged. Moreover the application of multidimensional models for the residuals in the case of detected connection, can be prohibitive due to the dimension of the problem if it is done without simplification. The time series considered regard the European Stock Indexes monthly figures from 1 January 1995 to 31 December 2012 taken by the Eurostat database.

## 2 Model Summary

A multivariate time series of the log return for the stock index $i$ at time $t$ is $r_{i,t}$ $i=1,...,k$ and $t=1,...,T$ , that's to say $r_{i,t}=log(p_{i,t}/p_{i,t-1})$ , $p_{i,t}$ is the index value at time $t$. We can then define the time series of the past $r_{i,t}^{lag1}=r_{i,t-1}$, $i=1,...,k$ and $t=1,...,T$ where $r_{i,0}=NA$ . A *graph* is defined as a couple $G=(V,E)$ where $V$ is a set of vertices or nodes and $E$ is a set of edges, [3]. Two vertices $\alpha$ and $\beta$ of $G$ are said *adjacent*, $\alpha\sim\beta$, if there is an edge between them. A *cycle* is a sequence of nodes $v_1,v_2..,v_{n-1},v_1$ of length $n$ where $v_i\sim v_{i+1}$. If a cycle has adjacent elements $v_i\sim v_j$ and $j\notin\{i-1,i+1\}$ then it is said to have a *chord*. If it has no cords it is said *chordless*. A graph with no cordless cycles of length $\geq$ 4 is called *triangulated*. A generalization of DAGs and undirected graphs are the *chain graphs*. Chain Graphs are graphs that have undirected edges or directed edges associated with their pair of nodes, subject to the constraint that no cycle containing directed edge(s) is permissible [5]. It can be seen that such graphs have a block representation which separates vertices into blocks such that inter-block edges are arrows, intra-block edges are undirected and the blocks are completely ordered. Let $A,B,C$ be subset of $V$ and $f$ a joint density function of $(X_v)_{v\in V}$ then $A\perp B|C$ if

$f(x_A,x_B|x_C) = f(x_A|x_C)f(x_B|x_C)$ for each possible value $x_A$ of $X_A$,$x_B$ of $X_B$,$x_C$ of $X_C$. In the hypothesis of multivariate gaussian variables , $f = N(\mu,\Sigma), K = \Sigma^{-1}$, if $V$ represents the set of this variables, it is possible to define a dependence graph $G_K = (V,E_K)$ with $k_{u,v} = 0$ whenever there is no edge between a pair of vertices $u,v \in V$ .So if two vertices $u,v \in V$ are not adjacent, that's to say there's no edge, it holds that $u \perp v|V \setminus \{u,v\}$, $u$ and $v$ are *conditionally independent* given the others . Firstly $V$ contains the set of variables $Y$ given by $r_i, i = 1,...,k$ followed by $r_i^{lag1}, i = 1,...,k$ with $r_i = (r_{1,1},...,r_{1,T})$, $r_i^{lag1} = (r_{1,1}^{lag1},...,r_{1,T}^{lag1})$, $f = N(\mu,\Sigma)$. The method used to select a chain graph, that can deal with high dimension, is to search among the decomposable graph. *Decomposable* models are graphs that are triangulated. For this type of graphs a closed-form for the maximum likelihood estimates are available allowing computational simplification. The starting graph is the minimal AIC or BIC forest. A *forest* is an acyclic undirected graph, that is, an undirected graph with no cycle. Then search the edge giving the maximal AIC or BIC reduction is added until further reduction are not possible. Here the criterion is $-2logL + kdim(G)$, where $dim(G)$ is the number of edges of $G$, $L$ is the likelihood, $k = 2$ (AIC), $k = log(n)$ (BIC). Then it is possible to split the set of variables $\Omega_{t-1} = \{r_i, i = 1,...,k\}$ in two parts. A subset of $m < k$ variables $C = \{r_i, i = i_1,...,i_m\}$ which are connected to some element of $\Omega_{t-1} = \{r_i^{lag1}, i = 1,...,k\}$, with the exception of their own past. A subset of $k-m$ variables $\overline{C} = \{r_j, j = j_1,...,j_{k-m}\}$ which are not connected to any element of $\Omega_{t-1} = \{r_i^{lag1}, i = 1,...,k\}$, with the exception of their own past. For the variables in $\overline{C}$ we impose the link with their own past, to treat them all together with the variables in $C$ in a multiple restricted regression and compare with the full counterpart. That is called an adjustment.

As the concurrent correlations cannot be used in the forecast [4], only the past to present information is adjusted and used. Moreover as for each variable its equation in the model *RVAR* is a multiple normal regression the following result, which links regression and graph theory, is used:the $i$-th regression coefficient is 0 if only if the dependent variable is conditional independent from the $i$-th regression variable given the remaining variables,[5]. So The following model is estimated:

$$\begin{bmatrix} r_{1,t} \\ \vdots \\ r_{m,t} \end{bmatrix} = \begin{bmatrix} \phi_{0,1} \\ \vdots \\ \phi_{0,1} \end{bmatrix} + \begin{bmatrix} \phi_{1,1} & \cdots & \phi_{1,m} \\ \vdots & \ddots & \vdots \\ \phi_{m,1} & \cdots & \phi_{m,m} \end{bmatrix} \cdot \begin{bmatrix} r_{1,t-1} \\ \vdots \\ r_{m,t-1} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{m,t} \end{bmatrix} \tag{1}$$

where $a_{j,t}$ are gaussian errors, $\phi_{i,j}$ are parameters. The elements $\phi_{i,j}$ are restricted to 0 whenever no connection was founded between pair of variables and the past with the chain graph approach. This estimate has smaller asymptotic variance than the unrestricted estimator. As ARCH tests on the residuals indicate that data exhibit volatility cluster, that is, variance may be high for certain time periods and low for other periods, and variance jump are rare, for $\hat{a}_i$, a $MAarch(1)$ model is needed:

$$\begin{bmatrix} a^2_{1,t} \\ \vdots \\ a^2_{m,t} \end{bmatrix} = \begin{bmatrix} \psi_{0,1} \\ \vdots \\ \psi_{0,1} \end{bmatrix} + \begin{bmatrix} \psi_{1,1} & \cdots & \psi_{1,m} \\ \vdots & \ddots & \vdots \\ \psi_{m,1} & \cdots & \psi_{m,m} \end{bmatrix} \cdot \begin{bmatrix} a^2_{1,t-1} \\ \vdots \\ a^2_{m,t-1} \end{bmatrix} + \begin{bmatrix} b_{1,t} \\ \vdots \\ b_{m,t} \end{bmatrix} \tag{2}$$

where $\psi_{i,j}$ are parameters, $b_{j,t}$ is consider like a error. The elements $\psi_{i,j}$ are restricted to 0 whenever no connection was founded between pair of squared residuals and their past using a second chain graph approach as before. This estimate has smaller asymptotic variance than the unrestricted estimator.

## 3 Results

The time series considered regard the European Stock Indexes monthly figures from 1 January 1995 to 31 December 2012 taken by the Eurostat database. The following figure represents the Chain Graph for the 31 Indexes and their past, **fig. 1** $RVAR(1)$ part, and the successive Chain Graph for their squared RVAR residuals and their past, **fig. 2** $MArch(1)$ part, using a forward search among the decomposable graph minimizing the *BIC* criteria. Please note that for the figure 1,2 the left side is the present slide of the set of variables considered, whereas the right side is the past slide of the set of variables considered (lag1 is attached at the end of their names). So the edges between should be considered as directed from present to past. Moreover All the models are stable. It can be seen from the figures the parameters saving: 960 parameters are needed with respect to a full VAR, whereas 39 parameters are needed with respect to a Restricted VAR. Moreover 954 parameters are needed with respect to a full MArch, whereas 44 parameters are needed with respect to a Restricted MArch. The *AIC* for the full VAR is $-21237.63$ where for the Restricted VAR is $-21428.61$. So there is a slightly improvement by using the restricted model. The *AIC* for the full MArch is $-41513.95$ where for the Restricted March is $-41231.29$. So there is a slightly worsening by using the restricted model. It can be seen that the Full VAR back forecasts are near the Restricted VAR back forecasts and with some exceptions near to the actual values. In terms of mean squared error, [4],

$$MSE = \frac{\sum_{i\in\{1,\ldots,m\}} \sum_{t\in\{1,\ldots,T\}} (\hat{\delta}_{i,t} - \delta_{i,t})^2}{m \cdot T} \tag{3}$$

it results $MSE_{full,r} = 0.007975$ , $MSE_{rest,r} = 0.009584$ with $\delta_{i,t} = r_{i,t}$. So there is a slightly worsening by using the restricted model. However the parameters saving and the AIC improvement suggest to restrict the models. The equation (3) with with $\delta_{i,t} = \sigma_{i,t}$ gives $MSE_{full,\sigma} = 0.01045866$ , $MSE_{rest,\sigma} = 0.009379994$. The rolling $\sigma_t$ are considered as the actual values. Now there is a slightly worsening by using the full model. The multivariate portmanteau test applied to standardized squared residuals $\hat{a}^2_t/\hat{\sigma}^2_t$ to check for adequacy of the Full VAR and the Restrected VAR with respect to the ARCH effect up to lag 3 gives statistics 2957.6964 and 3408.745

with 2883 df and p-values of 0.1625 and 0. In **tab. 1** it can be seen from the p-values $< 0.05$ that the majority of the RVAR parameters are significant.

## 4 Concluding remarks

The great reduction in parameters seems not to affect the predictive power of the model VAR and MArch models as the MSEs are similar. Even if conditional normal distributions are chosen for $a_t$, their unconditional distributions result non normal in both cases thanks to the MArch. Moreover, at the one hand the full model manage to take care of the ARCH effects whereas the restricted doesn't. At the other hand the AIC is better for the restricted model. So it seems that for large problem and forecast use the restricted model has to be chosen.

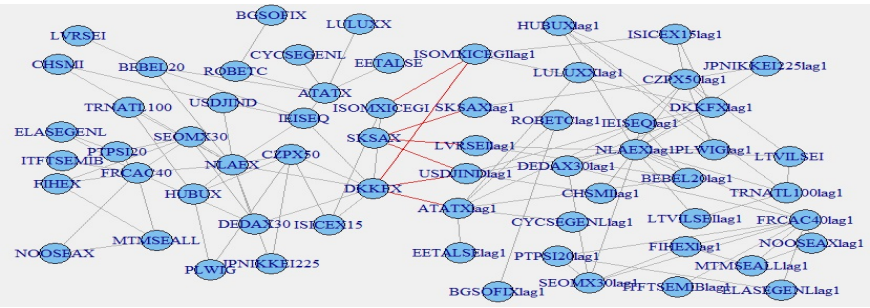


**Fig. 1** Chain Graph for $RVAR(1)$

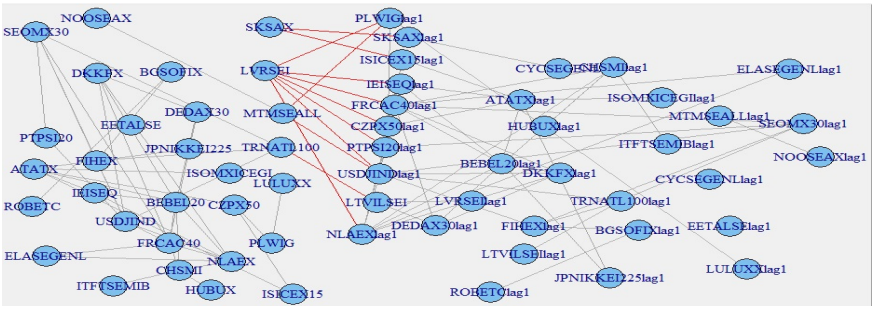


**Fig. 2** Chain Graph for $MARCH(1)$

| Variable$_t$ | Variables$_{t-1}$ | Estimate | Std.Error | tvalue | $Pr(>\|t\|)$ |
|---|---|---|---|---|---|
| BEBEL20$_t$ | BEBEL20$_{t-1}$ | 0.3160 | 0.0649 | 4.8636 | 0 |
| BGSOFIXB$_t$ | BGSOFIX$_{t-1}$ | 0.0566 | 0.0684 | 0.8285 | 0.4082 |
| CZPX50$_t$ | CZPX50$_{t-1}$ | 0.3632 | 0.0632 | 5.7382 | 0 |
| DKKFX$_t$ | DKKFX$_{t-1}$ | 0.08409 | 0.1012 | 0.8303 | 0.4072 |
| DKKFX$_t$ | ATATX$_{t-1}$ | 0.0362 | 0.0896 | 0.4043 | 0.6863 |
| DKKFX$_t$ | ISOMXICEGI$_{t-1}$ | 0.01848 | 0.0376 | 0.4910 | 0.6238 |
| DKKFX$_t$ | USDJIND$_{t-1}$ | 0.2767 | 0.1330 | 2.0807 | 0.0386 |
| DEDAX30$_t$ | DEDAX30$_{t-1}$ | 0.3163 | 0.0649 | 4.8676 | 0 |
| EETALSE$_t$ | EETALSE$_{t-1}$ | 0.1328 | 0.0678 | 1.9583 | 0.0514 |
| ELASEGENL$_t$ | ELASEGENL$_{t-1}$ | 0.3787 | 0.0633 | 5.9744 | 0 |
| FRCAC40$_t$ | FRCAC40$_{t-1}$ | 0.2897 | 0.0655 | 4.4177 | 0 |
| IEISEQ$_t$ | IEISEQ$_{t-1}$ | 0.3422 | 0.0643 | 5.3165 | 0 |
| ITFTSEMIB$_t$ | ITFTSEMIB$_{t-1}$ | 0.1801 | 0.0668 | 2.6968 | 0.007 |
| CYCSEGENL$_t$ | CYCSEGENL$_{t-1}$ | 0.2800 | 0.0657 | 4.2582 | 0 |
| LVRSEI$_t$ | LVRSEI$_{t-1}$ | 0.0297 | 0.0684 | 0.4341 | 0.6646 |
| LTVILSEI$_t$ | DEDAX30$_{t-1}$ | 0.0547 | 0.1843 | 0.2967 | 0.7669 |
| LTVILSEI$_t$ | LVRSEI$_{t-1}$ | 0.1071 | 0.1278 | 0.8380 | 0.4029 |
| LTVILSEI$_t$ | LTVILSEI$_{t-1}$ | 0.0811 | 0.0692 | 1.1714 | 0.2427 |
| LULUXX$_t$ | LULUXX$_{t-1}$ | 0.0349 | 0.0684 | 0.5100 | 0.6105 |
| HUBUX$_t$ | HUBUX$_{t-1}$ | 0.3516 | 0.0636 | 5.5233 | 0 |
| MTMSEALL$_t$ | MTMSEALL$_{t-1}$ | 0.1385 | 0.0678 | 2.0417 | 0.0424 |
| NLAEX$_t$ | NLAEX$_{t-1}$ | 0.2933 | 0.0655 | 4.4785 | 0 |
| ATATX$_t$ | ATATX$_{t-1}$ | 0.4182 | 0.0622 | 6.7223 | 0 |
| PLWIG$_t$ | PLWIG$_{t-1}$ | 0.3036 | 0.0647 | 4.6907 | 0 |
| PTPSI20$_t$ | PTPSI20$_{t-1}$ | 0.3671 | 0.0637 | 5.7615 | 0 |
| ROBETC$_t$ | ROBETC$_{t-1}$ | 0.0176 | 0.0684 | 0.2577 | 0.7968 |
| SKSAX$_t$ | LVRSEI$_{t-1}$ | 0.0769 | 0.0440 | 1.7458 | 0 |
| SKSAX$_t$ | SKSAX$_{t-1}$ | 0.3386 | 0.064 | 5.2617 | 0 |
| SKSAX$_t$ | USDJIND$_{t-1}$ | -0.0697 | 0.0925 | -0.7533 | 0 |
| SEOMX30$_t$ | SEOMX30$_{t-1}$ | 0.3464 | 0.064 | 5.3903 | 0 |
| FIHEX$_t$ | FIHEX$_{t-1}$ | 0.3822 | 0.0632 | 6.0420 | 0 |
| TRNATL100$_t$ | TRNATL100$_{t-1}$ | 0.3471 | 0.0642 | 5.4013 | 0 |
| CHSMI$_t$ | CHSMI$_{t-1}$ | 0.2877 | 0.0656 | 4.3856 | 0 |
| ISICEX15$_t$ | ISICEX15$_{t-1}$ | 0.1532 | 0.0677 | 2.2636 | 0.0246 |
| ISOMXICEGI$_t$ | ISOMXICEGI$_{t-1}$ | 0.4425 | 0.0614 | 7.2024 | 0 |
| NOOSEAX$_t$ | NOOSEAX$_{t-1}$ | 0.0335 | 0.0684 | 0.4896 | 0.6249 |
| JPNIKKEI225$_t$ | JPNIKKEI225$_{t-1}$ | 0.2553 | 0.066 | 3.8622 | 0.0001 |
| USDJIND$_t$ | USDJIND$_{t-1}$ | 0.2276 | 0.066 | 3.4083 | 0.0007 |

**Tab. 1:** $\hat{\phi}_{i,j}$ RVAR estimation

# References


1. Bollerslev T., Engle R.F., Nelson D.B., ARCH Model, Handbook of Econometrics, Elsevier Science, Amsterdam, vol. IV, pp. 2959 3038, (1994).
2. Campbell, J.Y., Chan Y.L., Viceira L.M., A Multivariate Model of Strategic Asset Allocation, Journal of Financial Economics, vol. 67, n.1, pp. 41 80, (2003).
3. Hojsgaard S., Edwards D.,Lauritzen S., Graphical Models with R, Springer, (2012).
4. Tsay R., Analysis of Financial Time Series, Wiley and Sons, (2010).
5. Whittaker J., Graphical Models in Applied Multivariate Statistics, Wiley and Sons, (1990).