“Dynamic dependence structure between energy markets and the Italian stock index”

**AUTHORS**
Giovanni Masala [https://orcid.org/0000-0003-1719-641X](https://orcid.org/0000-0003-1719-641X)

**ARTICLE INFO**

**DOI**
http://dx.doi.org/10.21511/imfi.15(2).2018.06

**RELEASED ON**
Thursday, 03 May 2018

**RECEIVED ON**
Sunday, 25 February 2018

**ACCEPTED ON**
Tuesday, 17 April 2018

**LICENSE**
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

**JOURNAL**
"Investment Management and Financial Innovations"

**ISSN PRINT**
1810-4967

**ISSN ONLINE**
1812-9358

**PUBLISHER**
LLC “Consulting Publishing Company “Business Perspectives”

**FOUNDER**
LLC “Consulting Publishing Company “Business Perspectives”

**NUMBER OF REFERENCES**
21

**NUMBER OF FIGURES**
3

**NUMBER OF TABLES**
5

© The author(s) 2018. This publication is an open access article.
The dependence structure between the main energy markets (such as crude oil, natural gas, and coal) and the main stock index plays a crucial role in the economy of a given country. As the dependence structure between these series is dramatically complex and it appears to change over time, time-varying dependence structure given by a class of dynamic copulas is taken into account.

To this end, each pair of time series returns with a dynamic t-Student copula is modelled, which takes as input the time-varying correlation. The correlation evolves with the DCC(1,1) equation developed by Engle.

The model is tested through a simulation by employing empirical data issued from the Italian Stock Market and the main connected energy markets. The author considers empirical distributions for each marginal series returns in order to focus on the dependence structure. The model’s parameters are estimated by maximization of the log-likelihood. Also evidence is found that the proposed model fits correctly, for each pair of series, the left tail dependence coefficient and it is then compared with a static copula dependence structure which clearly underperforms the number of joint extreme values at a given confidence level.

**Keywords**
- dependence structure, dynamic copulas, DCC model, tail dependence, energy markets, MIB stock market

**JEL Classification**
- C15, C63, G17

**INTRODUCTION**

The dependence structure between the energy markets and the main stock index plays a fundamental role in the strategic development of the economy of a given country. Particularly, the link between the main energy markets and their influence on the financial market dynamics for a given country should be analyzed. A comprehensive survey of this feature has been carried out by Gatfaoui (2016). In this paper, the author investigates the links between the U.S. natural gas and U.S. crude oil markets, as well as their dependencies with the U.S. stock market (Standard and Poor’s 500 Index).

At first, Gatfaoui performs a structural break test in order to detect eventual regime changes concerning the energy spot prices and stock market index over the given sample horizon. The second stage foresees to use copula functions, within each time segment, to fit the dependence structure involving the given energy commodities and the U.S. stock market.

Some authors such as Adams et al. (2017), Engle and Sheppard (2001) and Tse (2000) developed several tests to ascertain the variability of correlation. From these surveys, it appears that the correlations are generally time varying in the context of financial markets.
The dynamic copulas framework allows to effectively handle a time-varying dependence structure. We apply then a DCC(1,1) model introduced by Engle (2002). An application concerning basket cliquet options has been developed by Masala (2013).

To our knowledge, the literature did not investigate yet the dependence structure among energy markets and stock price index from a dynamic framework. In this survey, we investigate the Italian market and the main energy markets (crude oil, natural gas, and coal) to test the proposed model. We can further extend the same approach to any other markets.

The paper includes the following sections. The main goals of the survey are introduced in the current section. In section 1, we describe the theoretical backgrounds about the copula dependence structure in the static and dynamic case. Section 2 illustrates the characteristics and the main statistics about the databases used. We perform a numerical application in section 3 in order to apply the proposed models. Finally, last section concludes.

1. **THE DEPENDENCE STRUCTURE**

1.1. Copula functions

Copula functions allow to represent multivariate distributions through complex and non-linear dependence structures between its marginal distributions. Indeed, the copula splits a joint distribution into its marginals and the dependence amongst them. A great variety of copulas is available in the literature.

Abe Sklar (1959) introduced copula functions. Since 1986, the number of research papers about copula functions increased, starting from the seminal paper of Genest and MacKay (1986).

We also list actuarial applications by Frees and Valdez (1998) and financial applications by Embrechts et al. (2002).

**Definition 1.** A function $C$ is a two-dimensional copula if it matches:

i. domain $[0,1] \times [0,1]$;

ii. $C(0,u) = C(u,0) = 0$; $C(u,1) = C(1,u) = u$ for every $u \in [0,1]$;

iii. $C$ is 2-increasing:

$$C(v_1,v_2) + C(u_1,u_2) \geq C(v_1,u_2) + C(u_1,v_2)$$

for every $(u_1,u_2), (v_1,v_2) \in [0,1] \times [0,1]$ with $0 \leq u_1 \leq v_1 \leq 1$ and $0 \leq u_2 \leq v_2 \leq 1$.

The main result is the so-called fundamental theorem due to Sklar (1959).

**Theorem 1.** Given a two-dimensional distribution $F$ with marginals $F_1$ and $F_2$, we can find a two-dimensional copula $C$ with

$$F(x_1,x_2) = C(F_1(x_1), F_2(x_2)).$$

(1)

The continuity of $F_1$ and $F_2$ ensures the unicity of the copula $C$. This definition can be extended with higher dimensions.

1.2. Conditional copulas

Time series often contain random vectors with conditioning variables (represented as time functions $\Omega_t$). To face this situation, we introduce the so-called conditional copula (see Patton, 2006).

**Definition 2.** Conditional copulas in two dimensions.

A conditional copula $(x,y) | \Omega_t$, with $x | \Omega_{t-1} \sim F_t$ and $y | \Omega_{t-1} \sim G_t$, is represented by the bidimensional conditional distribution function of $U_t \sim F_{t}(x | \Omega_{t-1})$ and $V_t \sim G_{t}(y | \Omega_{t-1})$ given $\Omega_{t-1}$.

The Sklar theorem can be adapted as follows (see Patton, 2006).

**Theorem 2.** We denote $F_t$, $G_t$ and $H_t$ the conditional distributions of $x$, $y$ and $(x,y)$ respectively under $\Omega_{t-1}$ with $F_t$ and $G_t$ continuous in $x$ and $y$. We can prove the uniqueness of the copula $C_t$ satisfying:
\[ H_t(x,y|\Omega_{t-1}) = C_t\left(F_t(x|\Omega_{t-1}), G_t(y|\Omega_{t-1})|\Omega_{t-1}\right). \] (2)

Note that the converse is true.

We assume the same variable for the margins and the copula in order to ensure the validity of this result.

Some applications of dynamic copulas are given by Ausin and Lopes (2010), Fantazzini (2008), Masala (2013), and Manner and Reznikova (2011).

1.3. Marginal modeling

Remember that the joint density function writes as:
\[ h(x,y|\Omega_{t-1}; \vartheta_h) = f_t(x|\Omega_{t-1}; \vartheta_f) \cdot g_t(y|\Omega_{t-1}; \vartheta_g), \] (3)

where \( u \equiv F_t(x|\Omega_{t-1}; \vartheta_f) \), \( v \equiv G_t(y|\Omega_{t-1}; \vartheta_g) \), \( \vartheta_f \) and \( \vartheta_g \) represent the vectors of marginal parameters and \( \vartheta_h \) is the vector of copula parameters, with \( \vartheta_h = [\vartheta_f, \vartheta_g, \vartheta_c] \).

By taking logs in equation (1), we obtain the log-likelihood function:
\[ L_y(\vartheta_h) = L_x(\vartheta_f) + L_y(\vartheta_g) + L_c(\vartheta_f, \vartheta_g, \vartheta_c) \] (4)

with:
\[
\begin{align*}
L_{xy}(\vartheta_h) &= \log h(x,y|\Omega_{t-1}; \vartheta_h) \\
L_x(\vartheta_f) &= \log f_t(x|\Omega_{t-1}; \vartheta_f) \\
L_y(\vartheta_g) &= \log g_t(y|\Omega_{t-1}; \vartheta_g) \\
L_c(\vartheta_f, \vartheta_g, \vartheta_c) &= \log c_t(u,v|\Omega_{t-1}; \vartheta_c)
\end{align*}
\] (5)

We estimate the parameters with the Inference Functions for Margins (IFM) procedure, as listed below.

1. At first, determine the parameters \( \vartheta_f \) and \( \vartheta_g \) of the marginals \( F_t \) and \( G_t \) by maximizing the log-likelihood:
\[ \hat{\vartheta}_f = \text{arg max } L(\vartheta_f) = \text{arg max } \sum_{t=1}^{T} \log f_t(x_t; \vartheta_f) \] (6)

\[ \hat{\vartheta}_g = \text{arg max } L(\vartheta_g) = \text{arg max } \sum_{t=1}^{T} \log g_t(y_t; \vartheta_g) \] (7)

If the parameters are not ‘variation free’, we set:
\[ \{\hat{\vartheta}_f, \hat{\vartheta}_g\} = \text{arg max } \left[ L(\vartheta_f) + L(\vartheta_g) \right] = \text{arg max } \sum_{t=1}^{T} \left[ \log f_t(x_t; \vartheta_f) + \log g_t(y_t; \vartheta_g) \right]. \] (8)

2. Then use the estimations of the previous step to determine the copula’s parameters \( \vartheta_c \):
\[ \hat{\vartheta}_c = \text{arg max } L(\vartheta_c) = \text{arg max } \sum_{t=1}^{T} \log \left[ c_t(x_t; \vartheta_f), G_t(y_t; \vartheta_g); \vartheta_c \right] \] (9)

Note that the marginal distributions are fitted with the more suitable distributions. Nevertheless, in order to focus on the dependence structure, we can use an empirical distribution such as in Gatfaoui (2016).

See also Géégan and Zhang (2008), Hafner and Reznikova (2010), and Palaro and Hotta (2006) for a similar approach.

Remember that the correlation \( \rho_t \) follows the same equation as the DCC(1,1) structure established by Engle (2002):
\[ \rho_t = (1 - \lambda - \gamma) \cdot \bar{Q} + \lambda \cdot \varepsilon_{t-1} + \gamma \cdot \rho_{t-1} \] (10)

Here, \( \bar{Q} \) denotes the sample correlation and \( \varepsilon_{t-1} \) are the transformed standardized residuals and the parameters must satisfy the condition:
\[ \lambda + \gamma < 1 \quad \lambda, \gamma \in (0,1). \] (11)

This model was also applied by Patton (2006), Jondeau and Rockinger (2006) and Dias and Embrechts (2010). The copulas estimations were performed with a Matlab toolbox, described in Vogiatzoglou (2010).

Finally, we simulate trajectories for each pair of variables through a Monte Carlo procedure.

We list hereafter the steps of the algorithm (see Fantazzini, 2008):
• set the starting correlation value, next values will then satisfy the relation 
\[ \rho_t = \left( 1 - \lambda - \gamma \right) \cdot Q_t + \lambda \cdot \Psi_t \cdot s - \gamma \cdot \rho_{t-1}; \]

• denote \( \Lambda_t \), the Cholesky decomposition of the correlation matrix \( \Sigma_t; \)

• generate two independent values \( z = (z_1, z_2) \) from the \( N(0,1) \) distribution;

• generate a value \( s \) from \( \chi^2_\nu \) distribution (independent from \( z \));

• estimate the vector \( y_t = \Lambda_t \cdot z; \)

• set \( x_t = \frac{\sqrt{D}}{s} \cdot y_t; \)

• get the components \( R_t(x_i) = t_{\nu}(x_{i,t}) \) with \( i = 1, 2 \). The final vector is then \( (R_1(t), R_2(t))^T \sim C^2_\nu, \Sigma \).

2. THE DATABASE DESCRIPTION

Historical time series for oil (Europe Brent Spot Price FOB – dollars per barrel) and gas (Henry Hub Natural Gas Spot Price – dollars per million Btu) prices are available on the U.S. Energy Information Administration page [http://www.eia.gov](http://www.eia.gov). Specifically, the sources for these data are respectively:

- [http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRTE&f=D](http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RBRTE&f=D) (available from May 20, 1987);


We exhibit in Table 1 the descriptive statistics of these data and their returns for the period April 2, 2008–December 29, 2017.

We then exhibit in Figure 1 the scatter plots of the historical returns between each pair of variables.

We have already noted that correlation is generally time-varying. To verify this assumption, we plot in Figure 2 the correlation for each pair of variables’ returns for each year in the period 2004–2017 (note that year 2008 is not full).

From these figures, it is clear that correlations between each pair of returns variables are fluctuating over time, therefore a time-varying correlation model, like the employed copula – DCC model is a suitable choice.

The correlation matrix between the return series (for the whole database) is given in Table 2.

### Table 1. Descriptive statistics: MIB, oil, gas, coal and associated returns

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Prices MIB Oil Gas Coal</th>
<th>Returns MIB Oil Gas Coal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>19,869.12</td>
<td>81.30</td>
</tr>
<tr>
<td>Standard dev.</td>
<td>3,702.09</td>
<td>28.51</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.03</td>
<td>-0.01</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.44</td>
<td>-1.44</td>
</tr>
<tr>
<td>Maximum</td>
<td>34,547.00</td>
<td>143.95</td>
</tr>
<tr>
<td>Minimum</td>
<td>12,362.51</td>
<td>26.01</td>
</tr>
<tr>
<td>Count</td>
<td>2,424</td>
<td>2,424</td>
</tr>
</tbody>
</table>
Another key point concerns the presence of tail dependence between each pair of variables $X$ and $Y$. For example, left tail dependence is a fundamental concern for risk theory. For this purpose, we should estimate the following conditional probabilities:

$$P\left\{ X < X^* \mid Y < Y^* \right\},$$  \hspace{1cm} (12)

where $X^*$ and $Y^*$ are fixed values for each pair of variables. We select these thresholds as corresponding to a fixed quantile of the variables and compare the subsequent conditional probability with the independence or other conditions on the tails.

Let us examine the procedure established by van Oordt and Zhou (2012). The tail parameter for two variates $X$ and $Y$ is:

$$\tau_{Y|X} = \lim_{p \to 0} \frac{\Pr\{ Y < Q_Y(p) \land X < Q_X(p) \}}{p},$$  \hspace{1cm} (13)

where $Q_Y(p)$ denotes the quantile of the distribution of $Y$ at probability level $p$. 

![Figure 1. Scatter plot for each pair of returns](image1)

![Figure 2. Linear correlation coefficients between each pair of returns variables: MIB (1), oil (2), gas (3) and coal (4).](image2)
The non-parametric estimator of the $\tau_{Y\mid X}$ measure is the ratio between the number of observations in which both $X$ and $Y$ are extremes and those in which variable $X$ is extreme.

The non-parametric estimator of $\tau_{Y\mid X}$ is given by:

$$\hat{\tau}_{Y\mid X} = \frac{\sum_{t=1}^{\overline{n}} I_{Y_{x,t}}}{\sum_{t=1}^{\overline{n}} I_{X_{x,t}}},$$

(14)

where $I_{Y_{x,t}} = 1(Y_t < Q_Y(k/n) \wedge X_t < Q_X(k/n))$, $I_{X_{x,t}} = 1(X_t < Q_X(k/n))$ and $\overline{n}$ is the number of observations. We denote $1(\cdot)$ the indicator function, and we estimate the quantile function $Q(k/n)$, as usual, by the $k$-th lowest observation.

We deduce that the non-parametric estimator coincides with the OLS estimate of the slope coefficient as we perform a regression of the tail values of $Y$ against the indicator for tail values of $X$ (with no the constant term). Namely, the OLS estimate of $\beta$ in the indicator regression:

$$I_{Y_{x,t}} = \beta \cdot I_{X_{x,t}} + \epsilon_t.$$  

(15)

Note that the right tail is defined analogously. We just have to adapt the following quantities:

$$I_{Y_{x,t}} = 1(Y_t > Q_Y(k/n) \wedge X_t > Q_X(k/n))$$

(16)

and $I_{X_{x,t}} = 1(X_t > Q_X(k/n))$.

We exhibit in the next Table 3 the estimation $\hat{\tau}_{Y\mid X}$ at 95% level for each pair of variables.

3. THE NUMERICAL APPLICATION

We aim to replicate, in this section, each pair of returns series. The general framework foresees the following steps:

- estimation for the two marginals;
- parameters’ estimation of the dynamic copula function;
- simulation of the bivariate series.

3.1. Marginal distributions

As stated before, we adopt a non-parametric model for the marginals. To this end, we consider the empirical distribution function. More precisely, we approximate it with a smoothing

<table>
<thead>
<tr>
<th>Test statistics</th>
<th>MIB/oil</th>
<th>MIB/gas</th>
<th>MIB/coal</th>
<th>Oil/gas</th>
<th>Oil/coal</th>
<th>Gas/coal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>7.2214</td>
<td>31.2342</td>
<td>5.9513</td>
<td>16.1369</td>
<td>4.6686</td>
<td>26.3003</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.0053</td>
<td>0.0105</td>
<td>0.0060</td>
<td>0.0016</td>
<td>0.0009</td>
<td>0.0022</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.9940</td>
<td>0.9124</td>
<td>0.9924</td>
<td>0.9979</td>
<td>0.9990</td>
<td>0.8344</td>
</tr>
<tr>
<td>BIC</td>
<td>–375.1324</td>
<td>17.7166</td>
<td>–412.1775</td>
<td>–23.4716</td>
<td>–347.5480</td>
<td>5.1733</td>
</tr>
<tr>
<td>LogL</td>
<td>199.2553</td>
<td>2.8308</td>
<td>217.7779</td>
<td>23.4249</td>
<td>185.4631</td>
<td>9.1025</td>
</tr>
</tbody>
</table>

Table 3. Left tail parameter estimation for each pair of variables (number of cases between parentheses) at 95% level

Table 4. T-copula and DCC parameters for each pair of returns series
kernel function (performed with ‘ksdensity’ function, Matlab R2017a).

3.2. Copula parameters

The next step is to determine the dynamic dependence structure for each pair of variables. For this purpose, we consider a $t$-Student copula with parameter $\rho$, following the DCC model

$$Q_t = \rho_t \cdot \Psi_{t-1} + \gamma \cdot \rho_{t-1},$$

where $\rho_t$ is the parameter to be estimated, $\Psi_{t-1}$ is the previous correlation matrix, and $\gamma$ is the degree of freedom.

We reveal in Table 4 the parameters estimation and the test statistics for each pair of returns series.

Then, we plot in Figure 3 the dynamic correlation following the DCC equation for each pair of variables.

3.3. Monte Carlo simulation

The final step is to simulate bivariate values for each pair of returns with a Monte Carlo simulation. We apply so the algorithm described in section 2.3.

We then estimate the left tail coefficient for the three pairs that present high tail dependence (namely MIB/oil, MIB/coal and oil/coal). The results (average values obtained from Monte Carlo simulation) are unveiled in Table 5. We have added to this table the correlation coefficient.

Table 5. Left tail parameter estimation and correlation for some pairs of simulated variables (number of cases between parentheses)

<table>
<thead>
<tr>
<th>Pair</th>
<th>Left tail</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIB/oil</td>
<td>0.2278 (27.6)</td>
<td>34.16%</td>
</tr>
<tr>
<td>MIB/coal</td>
<td>0.3125 (37.8)</td>
<td>34.58%</td>
</tr>
<tr>
<td>oil/coal</td>
<td>0.2436 (29.5)</td>
<td>31.89%</td>
</tr>
</tbody>
</table>

These results, compared with Table 3, highlight that extreme joint values on the left tail of the bivariate distribution are suitably represented by the simulated values. Besides, the correlation coefficients are very similar with the real correlations given in Table 2.

CONCLUSION

In this survey, we developed a procedure able to determine the complex and dynamic dependence structure between a stock market index (MIB index for the Italian market) and the main energy markets (namely crude oil, natural gas and coal). A statistical inspection of the data emphasized that the correlations between each pair of series are far from being constant and they may assume sizeable values.
The scheme proposed in our survey assumes that a general class of dynamic copulas represent the dependence structure between each pair of variables to face the empirical nature of time-varying dependence structure. Besides, we simulate bivariate trajectories for the selected pair with Monte Carlo procedure. The marginal variables have been modeled through empirical distributions in order to focus only on the dependence structure.

The main issue is then to prove that the dynamic copula allows modelling faithfully the time-varying dependence structure. Besides, we have implemented the algorithms involved in order to replicate random bivariate series through Monte Carlo simulation. We have shown, for this purpose, that the left tail coefficients and the correlation are rather accurate for the simulated bivariate series.

REFERENCES


