NON–LINEAR MODELLIZATION OF BIVARIATE ASSET PRICES COMOVEMENTS AND APPLICATIONS

Marco Corazza[1], Anastasios G. Malliaris[2] and Elisa Scalco[1]

[1] Dipartimento di Matematica Applicata
Università Ca' Foscari di Venezia
(e-mail: corazza@unive.it, scalco@unive.it)

[2] Department of Economics
Loyola University of Chicago
(e-mail: tmallia@luc.edu)

ABSTRACT: We propose an investigating approach, and some related results and applications, for the non–linear modelling of a widely investigated phenomenon: the bivariate comovements between prices of assets (see, for example, MALLIARIS, A.G., & URRUTIA, J.L., 1996).

KEYWORDS: Bivariate dependence, asset prices, nonlinear modellization, energy asset, European call and put options, cross-Greeks.

1 Introduction

We propose an investigating methodology for the phenomenon of the nonlinear comovements among two asset prices; when such a bivariate dependence relationship is detected, our approach provides a polynomial approximation of it (section **2**). Then, we apply our investigating methodology to the prices time series of some energy assets (crude oil, gasoline and heating oil) traded in the U.S.A. (section **3**). Finally, we utilize this bivariate dependence relationship to achieve analytical approximations of the cross–Greeks of the vanilla European call and put options in terms of an asset whose price comoves with the price of the underlying of the investigated option (section **4**).

2 Our methodology

Our methodology is articulated in three steps:
• in the first step we propose a simple index able to evaluate the degree of bivariate dependence, and we provide some theoretical results about it;
• in the second step we propose a procedure by which to test the statistical meaningfulness of this index;
• once the meaningfulness of the index has been proved, in the third step we propose an algorithm which provides a polynomial approximation of the bivariate dependence relationship.

As far the first step is concerned, we start by considering two time series, $\{X_1(t), t = t_1, \ldots, t_N\}$ and $\{X_2(t), t = t_1, \ldots, t_N\}$. The simple index we propose is defined as follows:

Random variable	$\delta_{i,j}$	Bilateral $t-$test	Check on the $\delta_{1,2}-$dependence
$X_{CO}(t)$, $X_G(t)$	0.35407	R	P
$X_{CO}(t)$, $X_{HO}(t)$	0.39259	R	P
$X_G(t)$, X_{HO}	0.48642	R	P

Table 1.

$$\delta_{1,2} = \frac{1}{N-1}\sum_{t=t_2}^{t_N} \Delta(t)_{1,2},\ \Delta(t)_{1,2} = \begin{cases} -1 \text{ if } [X_1(t) - X_1(t-1)][X_2(t) - X_2(t-1)] < 0 \\ 1 \text{ if } [X_1(t) - X_1(t-1)][X_2(t) - X_2(t-1)] \geq 0 \end{cases}.$$

With regards to theoretical links between $\delta_{1,2}$ and the Bravais–Pearson linear correlation coefficient $\rho_{1,2}$, we give the following proposition:

Proposition 1. *Let $f(\cdot)$: $\mathbb{R}_+ \to \mathbb{R}_+$ be the bivariate dependence relationship between $X_1(t)$ and $X_2(t)$, $X_1(t) = f(X_2(t)) + \varepsilon(t)$, where $\varepsilon(t)$ has the usual meaning, and let $f(\cdot)$ be infinite times derivable in $m_2 = \mathbb{E}(X_2(t))$. If*

$$\frac{f^{(i)}(m_2)}{i!}(-m_2)^{i-j} = 0\ \forall\ i, j\ :\ i = 0, \ldots, +\infty \wedge j = 2, \ldots, +\infty \wedge i - j \geq 2,$$

where $f^{(i)}(\cdot)$ indicates the $i-$th derivative of $f(\cdot)$, then the bivariate dependence relationship is affine.

As far the second step is concerned, the philosophy of the procedural approach we propose for testing the statistical meaningfulness of $\delta_{1,2}$ is similar to the one of the procedural approach proposed in KABOUDAN, M.A., 2000. In particular, it allows to perform bilateral and unilateral $t-$tests.

As far the third step is concerned, once the meaningfulness of the index has been proved, we model the bivariate dependence relationship $X_1(t) = f(X_2(t)) + \varepsilon(t)$. In particular, we search for a polynomial approximation of $f(\cdot)$, *i.e.*

$$f(X_2(t)) = \sum_{j=0}^{K} a_j X_2^j(t) + r(K),$$

where K is the truncation order of the involved Taylor's series, $a_j = \sum_{i=j}^{K} \frac{f^{(i)}(m_2)}{i!}\binom{i}{i-j}\cdot$ $\cdot(-m_2)^{i-j}$, and $r(K)$ is a suitable remainder function. Of course, in such an approach a crucial role is played by K. For detecting its "optimal" value, we propose an algorithm whose search procedure is based on a standard cross–validation technique, as suggested for empirical work in POGGIO, T., & SMALE, S., 2003.

3 Applications to energy asset prices time series

In general terms, for each application we act as follows:
• we consider the generic bivariate time series $\{(X_1(t),\ X_2(t)),\ \ t = t_1, \ldots, t_N\}$;
• we split the last 10% of the bivariate time series and utilize it as forecasting data set D_F at the end of the application for performing an out–of–sample check;
• we split the remaining 90% percent of the bivariate time series into the learning data set D_L (its first 70%) and the validation data set D_V (its last 30%);

Polynomial approximation
$\widehat{X}_{CO}(t) = 1.68731 + 31.36198 X_G(t)$
$\widehat{X}_{CO}(t) = -9.38380 + 97.98516 X_{HO}(t) - 116.50908 X_{HO}^2(t) + 63.03937 X_{HO}^3(t)$
$\widehat{X}_G(t) = 0.03803 + 0.02687 X_{CO}(t)$
$\widehat{X}_G(t) = 0.12943 + 0.79407 X_{HO}(t)$
$\widehat{X}_{HO}(t) = -3.80625 + 0.87743 X_{CO}(t) - 0.06879 X_{CO}^2(t) + 0.00240 X_{CO}^3(t) -$ $- 0.00003 X_{CO}^4(t)$
$\widehat{X}_{HO}(t) = -0.19987 + 2.37713 X_G(t) - 2.98411 X_G^2(t) + 1.89758 X_G^3(t)$

Table 2.

• we perform our methodology by using D_L and D_V.

As far the data are concerned, each univariate time series is constituted by $2,026$ daily spot closing prices of one of the following energy assets traded in U.S.A.: the crude oil ($X_{CO}(t)$), the gasoline ($X_G(t)$), and the heating oil ($X_{HO}(t)$). Such prices have been collected from January 3, 1994 to February 6, 2002.

The exposition of the results is organized in two tables and one figure.

As far **Table 1** is concerned:

• the third column provides the response of the independence hypothesis test: "*A*" or "*R*" for, respectively, the acceptance or the rejection;

• if the independence hypothesis is rejected, then the fourth column gives the response of the test whether the $\delta_{i,j}-$dependence, with $i,\ j \in \{CO, G, HO\}$ and $i \neq j$, is negative ("*N*") or positive ("*P*").

A few remarks about the results reported in **Table 1**:

• the fact that $\delta_{i,j}$ is significantly different from 0 for all the considered i and j indicates the existence of a bivariate dependence relationship between each possible pair of $X_i(t)$ and $X_j(t)$;

• the fact that $\delta_{i,j}$ is positive for all the considered i and j can be interpreted as an indicator of the positiveness of the dependence between each possible pair of $X_i(t)$ and $X_j(t)$.

In **Table 2** we report the polynomial approximations for each possible pair of $X_i(t)$ and $X_j(t)$. Notice that the fact that the degree of the polynomial approximation is greater than 1 in a significant percentage of cases confirms the presence of non–linearities in some of the investigated dependence relationships.

Finally, as representative applications, we apply the first and the second polynomial approximations reported in **Table 2** to the corresponding data set D_F (see **Fig. 4**).

4 Cross–Greeks

Before to present the analytical approximations of the cross–Greeks, we premise some notation: we denote the $l-$th derivative of $X_1(X_2(t))$ by $X_1^{(l)}(X_2(t))$; we denote the function of the cumulative probability distribution of a standard normally distributed random variable and its first derivative, respectively by $\Phi(x)$ and $\Phi^{(1)}(x)$.

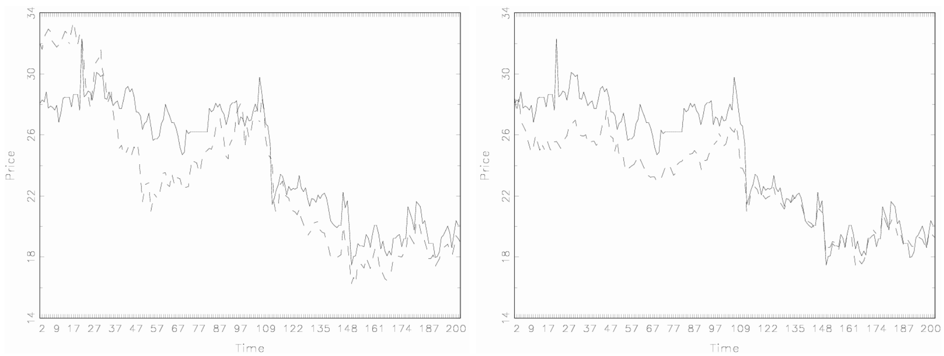

Figure 1. *On the right, the dotted uneven line represents the behaviour in D_F of the polynomial approximation of $X_{CO}(t)$ in terms of $X_G(t)$. On the left, the dotted uneven line represents the behaviour in D_F of the polynomial approximation of $X_{CO}(t)$ in terms of $X_{HO}(t)$.*

Proposition 2. *Let the usual hypotheses concerning the Black–and–Scholes environment hold, and let $X_1(t)$ and $X_2(t)$ be the prices of two assets, both defined on $[t_0, t_1]$ with $t_0 < t_1$. If $X_1(X_2(t)) = \sum_{i=0}^{K} a_i X_2^i(t)$, with $K \in \mathbb{N}^0$ and $a_i \in \mathbb{R}$, then*

$$
\begin{gathered}
cross-delta_{call} = \Phi(d_1^*) X_1^{(1)}(X_2(t)),\\
cross-gamma_{call} = \frac{\Phi^{(1)}(d_1^*)}{X_1(X_2(t))\sigma\sqrt{\tau}} \left[X_1^{(1)}(X_2(t))\right]^2 + \Phi(d_1^*) X_1^{(2)}(X_2(t)),\\
cross-vega_{call} = X_1(X_2(t))\sqrt{\tau}\,\Phi^{(1)}(d_1^*),\\
cross-theta_{call} = \frac{X_1(X_2(t))\sigma}{2\sqrt{\tau}}\Phi^{(1)}(d_1^*) + Xre^{-r\tau}\Phi(d_2^*),\\
cross-rho_{call} = X\tau e^{-r\tau}\Phi(d_2^*),\\
cross-delta_{put} = [\Phi(d_1^*) - 1] X_1^{(1)}(X_2(t)),\\
cross-gamma_{put} = \frac{\Phi^{(1)}(d_1^*)}{X_1(X_2(t))\sigma\sqrt{\tau}} \left[X_1^{(1)}(X_2(t))\right]^2 + [\Phi(d_1^*) - 1] X_1^{(2)}(X_2(t)),\\
cross-vega_{put} = X_1(X_2(t))\sqrt{\tau}\Phi^{(1)}(d_1^*),\\
cross-theta_{put} = \frac{X_1(X_2(t))\sigma}{2\sqrt{\tau}}\Phi^{(1)}(d_1^*) + Xre^{-r\tau}[\Phi(d_2^*) - 1] \text{ and}\\
cross-rho_{put} = X\tau e^{-r\tau}[\Phi(d_2^*) - 1],
\end{gathered}
$$

where $d_1^* = \frac{\log(X_1(X_2(t))/X) + r\tau + \sigma^2\tau/2}{\sigma\sqrt{\tau}}$ *and* $d_2^* = d_1^* - \sigma\sqrt{\tau}$.

References

KABOUDAN, M.A. 2000. Genetic programming prediction of stock prices. *Computational Economics*, **16**, 207–236.

MALLIARIS, A.G., & URRUTIA, J.L. 1996. Linkages between agricultural commodity futures contracts. *The Journal of Futures Markets*, **16**, 595–609.

POGGIO, T., & SMALE, S. 2003. The mathematics of learning: Dealing with data. *Notices of the American Mathematical Society*, **50**, 537–544.