Physica A 413 (2014) 284–293

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Cross-correlation between crude oil and refined product prices

Li Liu [a,1], Guofeng Ma [b,*]

[a] School of Finance, Nanjing Audit University, West Yushan Road 86, Pukou District, Nanjing, Jiangsu Province, China

[b] School of Economics and Management, Tongji University, Shanghai, China

HIGHLIGHTS

- This is the first paper on cross-correlations in energy markets.
- Cross-correlations between oil and refined product prices are significant.
- Cross-correlations are strong and display the property of multifractality.
- Conventional models have limited power in capturing long-range cross-correlations.
- Cross-correlations change over time.

ARTICLE INFO

Article history:
Received 3 March 2014
Received in revised form 6 June 2014
Available online 8 July 2014

Keywords:
Detrended cross-correlation analysis
Crude oil
Refined product
Multifractality

ABSTRACT

In this paper, we investigate cross-correlations between crude oil and refined product prices based on the well-known detrended cross-correlation analysis (DCCA). Our findings indicate that the cross-correlations are significant and strong. Furthermore, the multifractality in cross-correlations is also revealed. The cross-correlation coefficients are as high as 0.9 for larger time scales and are greater than those for smaller time scales. Two popular models, vector error correction model and bivariate BEKK volatility model, are found to have very limited power in capturing long-range cross-correlations, suggesting the drawbacks of these conventional models in actual applications. Long-term cross-correlations are stronger in recent ten years than those in the past decades.

1. Introduction

Crude oil can be refined into different components, named fractions (e.g., gasoline, heating oil and jet fuel), which can be used for many different purposes. Therefore, it is expected that crude oil and refined product prices are highly correlated. The plausible explanations are two aspects. First, crude oil is the major input of refining process. Higher oil prices imply greater refining cost and therefore lead to higher product prices. Second, if higher demand causes increases in refined product prices, refiners would like to produce more products to satisfy the demand. In this way, the increases (decreases) in oil input will cause higher (lower) crude oil prices.

There are a plethora of studies on the relationships between crude oil and refined product prices. The econometric methods used include the simple regression models [1–3], asymmetric vector error correction models (VECM) [4,5], threshold VECM [6,7], panel VECM [8], hidden cointegration [9] and the asymmetric autoregressive distributed lags (ARDL) [10]. A

* Corresponding author. Tel.: +86 15021858530.
E-mail addresses: liuli840821@126.com (L. Liu), gfmatju@126.com, mgfsjtu738@sina.com (G. Ma).

1 Tel.: +86 13611582191.

http://dx.doi.org/10.1016/j.physa.2014.07.007
0378-4371/

major drawback of these models is that they assume the linear relations between crude oil and refined product prices. It has been widely accepted that the joint dynamics between asset prices are nonlinear. Some econometric models allow for asymmetry and threshold adjustment of refined product prices and thereby seem to recognize the existence of nonlinear linkages. However, the asymmetry is actually an extension of linearity that the relations are linear in a specific regime but inconsistent between different regimes. Therefore, it is unreasonable to use abovementioned linear models to analyze the relationships between crude oil and refined product prices.

In this paper, we use a statistical test borrowed from statistical physics, named detrended cross-correlation analysis (DCCA) [11] to crude oil and refined product price data. This well-known method overcomes the drawback of conventional linear models by sufficiently allowing for the nonlinear joint dynamics and therefore is widely applied to financial data in existing studies [12–23]. Moreover, long-range cross-correlations revealed by DCCA are related to multivariate fractional Gaussian noise [24–26] which is more realistic than the classical assumption of multivariate Gaussian distribution in financial modeling [27]. Our findings based on DCCA can be summarized as follows.

First, as a preliminary analysis, we use a statistical test proposed by Podobnik et al. [28] and find that the cross-correlations are significant, especially for larger time scales. Second, the long-range cross-correlations are slightly anti-persistent and display strong multifractal behavior. Third, using a DCCA-based cross-correlation coefficient measure, we find that the cross-correlation coefficients for larger time scales are greater than those for smaller time scales. Fourth, as the work in most of existing studies, we use two seminal models, VECM for price relations and BEKK for volatility spillovers to capture joint dynamics between crude oil and refined products. Unfortunately, we find that these two models have very limited power in capturing long-range cross-correlations, further confirming the drawbacks of these conventional models in actual applications. Specifically, VECM can partly model cross-correlations at larger time scales whereas BEKK cannot capture cross-correlations at either smaller or larger time scales. Finally, we investigate whether cross-correlations change over time by analyzing cross-correlations in different subsample periods. Our evidence indicates that long-term cross-correlations are stronger in recent ten years than those in the past decades.

The remainder of this paper is organized as follows: Section 2 shows a description of DCCA. Section 3 gives data description and some preliminary analysis. Section 4 shows the empirical results. The last section concludes the paper.

2. Methodology

There are many methods explored to detect cross-correlations between financial time series such as cross-correlation function and the lagged random matrix theory [29]. However, these conventional methods cannot be applied to time series with nonstationarity or fat-tail distributions, both of which have been seminal stylized fact in financial time series. In this paper, we employ a detrended cross-correlation analysis (DCCA) [11] which is an extension of the well-known detrended fluctuation analysis [30]. The DCCA can well overcome abovementioned drawbacks of conventional methods and has been widely used to detect long-range cross-correlations in financial time series [12–23,31]. For two time series $\{x_t^*, t = 1, \ldots, T\}$ and $\{y_t^*, t = 1, \ldots, T\}$, where T is the length of data, the algorithm of DCCA contains following five steps.

First, we need to reorganize the time series by calculating the cumulative sums at each time and accordingly generate two new series, $x_k = \sum_{t=1}^{k}(x_t^* - \bar{x}^*)$ and $y_k = \sum_{t=1}^{k}(y_t^* - \bar{y}^*)$, $k = 1, \ldots, T$, where $\bar{x}^*$ and $\bar{y}^*$ are the average values of $\{x_t^*\}$ and $\{y_t^*\}$, respectively.

Second, we divide the two generated series x_k and y_k into $T_s = \text{int}(T/s)$ nonoverlapped subseries of the equal length of s. Since the length T of the series is often not a multiple of the considered time scale s, a short part at the end of each profile may remain. In order not to disregard this part of the series, the same procedure is repeated starting from the opposite end of each profile. Thereby, $2T_s$ segments are obtained together. We set $10 < s < N/5$.

Third, we calculate the co-moved covariance as,

$$F^2(s, \tau) \equiv \frac{1}{s}\sum_{j=1}^{s}[x_{(\tau-1)s+j} - \hat{x}_{(\tau-1)s+j}][y_{(\tau-1)s+j} - \hat{y}_{(\tau-1)s+j}] \tag{1}$$

for $\tau = 1, 2, \ldots, T_s$ and

$$F^2(s, \tau) \equiv \frac{1}{s}\sum_{j=1}^{s}[x_{T-(\tau-T_s)s+j} - \hat{x}_{T-(\tau-T_s)s+j}][y_{T-(\tau-T_s)s+j} - \hat{y}_{T-(\tau-T_s)s+j}] \tag{2}$$

for $\tau = T_s + 1, T_s + 2, \ldots, 2T_s$. The local trends $\hat{x}_{(\tau-1)s+j}$ and $\hat{y}_{(\tau-1)s+j}$ can be computed from linear, quadratic or higher order polynomial fit for each segment τ.

Fourth, we calculate the fluctuation function,

$$F(s) = \left\{\frac{1}{2T_s}\sum_{\tau=1}^{2T_s}[F^2(s, \tau)]\right\}^{1/2}. \tag{3}$$

Finally, we need to analyze the power-law relations between $F(s)$ and time scale s. If the power-law relations exist, we can say that the long-range cross-correlations exist, i.e.,

$$F(s) \sim s^{\lambda}. \tag{4}$$