<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
<th>DOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>201-231</td>
<td>Forecasting the density of returns in crude oil futures markets</td>
<td>Julien Chevallier</td>
<td>10.1504/IJGEI.2015.070264</td>
</tr>
<tr>
<td>242-256</td>
<td>Dynamic volatility spillover effect analysis between carbon market and crude oil market: a DCC-ICSS approach</td>
<td>Lean Yu; Jingjing Li; Ling Tang</td>
<td>10.1504/IJGEI.2015.070265</td>
</tr>
<tr>
<td>257-277</td>
<td>Spillover effect of international crude oil market on tanker market</td>
<td>Yuying Yang; Chang Liu; Xiaolei Sun; Jianping Li</td>
<td>10.1504/IJGEI.2015.070270</td>
</tr>
<tr>
<td>278-285</td>
<td>Oil price and economic growth: an improved asymmetric co-integration approach</td>
<td>Jian Chai; Ying Yang; Limin Xing</td>
<td>10.1504/IJGEI.2015.070269</td>
</tr>
<tr>
<td>286-297</td>
<td>Forecasting long-term and short-term crude oil price: a comparison of the predictive abilities of competing models</td>
<td>Zhongbao Zhou; Ke Duan; Ling Lin; Qianying Jin</td>
<td>10.1504/IJGEI.2015.070267</td>
</tr>
</tbody>
</table>
The dynamics of crude oil price movements: from financialisation of crude oil prices and price expectation perspectives
Youngho Chang; Shang Jyi Soon; Mirza Muhammad Hanif; Lucy Kusnadi
DOI: 10.1504/IJGEI.2015.070274

Oil demand forecasting for India using artificial neural network
S. Jebaraj; S. Iniyan
DOI: 10.1504/IJGEI.2015.070280

Road transport energy consumption in the G7 and BRICS: 1973-2010
Yi-Xuan Gao; Hua Liao; Paul J. Burke; Yi-Ming Wei
DOI: 10.1504/IJGEI.2015.070260

Additional Papers

A bibliometric analysis of energy poverty research: results from SCI-E/SSCI databases
Kang Li; Su-Yan Pan; Yi-Ming Wei
DOI: 10.1504/IJGEI.2015.070263

Long-term trends in non-renewable resource commodity prices: fresh evidence in the presence of structural breaks
Aviral Kumar Tiwari
DOI: 10.1504/IJGEI.2015.070271
Forecasting the density of returns in crude oil futures markets

Julien Chevallier

IPAG Business School,
IPAG Lab. 184,
Boulevard Saint-Germain,
75006 Paris, France
and
Université Paris 8 (LED),
2 rue de la Liberté,
93526 Saint-Denis, France
Email: julien.chevallier04@univ-paris8.fr

Abstract: Using tick-by-tick data for the WTI crude oil (2001–2010) market, this paper relies on the recent bivariate model by Maheu and McCurdy and compares the forecast accuracy of the density of returns with the HAR-RV model at different horizons up to 60 days. Our results provide evidence of the incremental information for density forecasting embedded in intraday data when the model is compared with the univariate EGARCH model. Turning to the comparison between the forecasting power of the realised volatility, which includes the total quadratic variation, vs. Bipower Variation (BPV) and median realised volatility, which only estimates the diffusive component, it is shown that the additional information contained in the jump component is significant on average. The findings for WTI crude oil futures confirm the importance of considering the continuous/jump decomposition for density forecasting.

Keywords: WTI crude oil; density forecasting; bivariate model.

Reference to this paper should be made as follows: Chevallier, J. (2015) ‘Forecasting the density of returns in crude oil futures markets’, Int. J. Global Energy Issues, Vol. 38, Nos. 4/5/6, pp.201–231.

Biographical notes: Julien Chevallier is an Associate Professor at the University Paris 8, as well as with the IPAG Business School (France). His interests lie in applied econometrics and empirical finance.

1 Introduction

Forecasting density is essential in empirical finance applications such as portfolio choice, risk management activities or derivatives pricing. Each activity requires indeed a full specification of the return distribution for the pricing of crude oil. While recent developments in financial econometrics allow to derive better forecasts of return densities (see Corradi and Swanson, 2006 for a recent survey), the issue of the inclusion of the jump component and its information content for such a purpose has not been documented to date for the crude oil market.
Diebold and Mariano (1995), West (1996) and White (2000) – among others – have pioneered the field of predictive density selection. In their context, accuracy is measured using a distributional analog of mean square error. Diebold and Mariano (1995) propose a well-known test for the null hypothesis of equal predictive ability that is based in part on the pairwise model comparison test discussed by Granger and Newbold (1986). Further on this issue, Giacomini (2002) suggests a weighted version of the Vuong (1989) likelihood ratio test for the case of dependent observations. Within the context of nested linear models, Clark and McCracken (2001, 2003) propose some easy to implement tests, under the assumption of martingale difference prediction errors. A limitation of the tests above is that they rule out possible dynamic misspecification under the null. A test which does not require correct dynamic specification and/or conditional homoskedasticity is proposed by Chao et al. (2001). Last but not least, we may refer to the ‘reality check’ approach by White (2000) and Hansen (2005) for a novel approach for dealing with the issue of choosing amongst many different models.

Hansen et al. (2012), Shephard and Sheppard (2010) and Maheu and McCurdy (2011) have suggested ‘complete’ models of returns and volatility. In particular, Maheu and McCurdy (2011) propose a bivariate specification of returns and volatility to obtain density forecasts at horizons up to 60 days. They confirm, in the density context, numerous previous findings that intraday data improve forecasts. The estimation of a multivariate models possesses interesting characteristics compared to the separate estimation of an univariate model. Namely, multivariate models allow to obtain densities forecasts. Maheu and McCurdy (2011) allow for a rich underlying distribution in the return equation, by using mixture of normals (see Bertholon et al., 2006). We merge these two strands of this recent literature to investigate whether the separation between the continuous and the jump components is of central importance in predicting the density of returns. Our results strongly argue in favour of separating the two components when forecasting the density of returns up to 60 days. Indeed, disentangling jumps from the continuous component help in forecasting the density of returns when applied to the case of WTI crude oil futures over 2001–2010.

This paper adopts the parsimonious specification of the Heterogeneous Autoregressive Model of the Realised Volatility (HAR-RV) model by Corsi (2009) to capture the well-known long-memory dependence in volatility. We also proceed with the detection of jumps following Huang and Tauchen’s (2005) statistical test relying on the Bipower Variation (BPV) estimator. We adapt the test statistic to the newly developed Median Realised Volatility (MedRV) following the empirical work by Theodossiou and Zikes (2009) showing the interesting properties of this estimator. The impact of microstructure noise for the WTI series is assessed statistically using the recent test by Awartani et al. (2009) in its jump-robust form. The bivariate model is estimated through maximum likelihood with possibly a mixture of normals, which allows to improve significantly the goodness-of-fit of the model.

This paper makes three contributions: (a) we extend the framework of Maheu and McCurdy (2011) and show how to model jumps in their bivariate framework; (b) we confirm their findings that intraday data yield to better densities forecasts than daily data for WTI futures and (c) we assess the importance of jumps when forecasting the density of returns by comparing jump-robust and non-jump-robust measures of realised volatilities. Compared to the ‘naive’ measure of realised volatility, considering jumps specifically provides significant improvement on the accuracy of forecasts of return
densities. We thus extend the results by Andersen et al. (2007a, 2007b) (see also Corsi et al., 2010) in showing the importance of disentangling jumps from the continuous component for forecasting purposes.

The remainder of the paper is organised as follows. Section 2 details our modelling strategy for jumps, the choice of volatility estimators using intraday data and the methodology to estimate the bivariate model. Section 3 discusses the empirical results. Section 4 concludes the paper.

2 Volatility, jumps and discrete time series model specification

In this section, we present first the time series models, second the parameter estimation procedures and third the density forecast comparison tests.

2.1 Time series models

The aim of our article is to present empirical evidence regarding the interest of disentangling jumps from volatility when it comes to forecasting the density of returns. To do so, we evaluate the relative forecasting performances of different discrete time series models. These models are selected for their ability to handle non-Gaussian distributions and time varying volatility. Jumps have indeed an effect on both the unconditional distribution of returns and volatility (Fang et al., 2012; Guégan et al., 2013). Their impact on times series models can be of three kinds:

- Jumps can be captured through the conditional distribution chosen in the discrete time series models.
- Jumps can impact the measurement of volatility, as revealed by the now large literature on realised risk measures.
- Jumps can impact the dynamics of volatility, as in most time series models the conditional volatility is computed from past returns or residuals that incorporate a jumps component by essence.

When a model would be able to handle these three aspects in a way that disentangles jumps from volatility components, the question of its superior ability to produce density forecast would still require to be considered. We base our empirical work on the three following time series models:

- A conditionally Gaussian model:
  \[ r_t = \mu + \sigma_t \varepsilon_t, \]  
  with \( \mu \) the returns expectation before jumps, \( \sigma_t \) the standard deviation, \( \varepsilon_t \sim N(0,1) \) the error term normally distributed. This model naturally ignores jumps. The only source of leptokurticity in the returns’ process comes from the time varying behavior of volatility.

- A model based on a mixture of two Gaussian distributions (MN) for its conditional distribution:
  \[ r_t = \mu + \sigma_t \varepsilon_t, \]  
  (2)
with $\varepsilon_t \sim MN(\theta, \mu_t, \sigma_t, \mu_x, \sigma_x)$. The subscripts 1 and 2 indicate mixing the two Gaussian distributions. Jumps are captured through $\varepsilon_t$ that is obtained by mixing two different Gaussian densities. As presented by Berthonol et al. (2006), this distribution is able to span a very large scope of couples of kurtosis and skewness. Possibly, this distribution is consistent with a mixture of a Gaussian distribution and of an extreme-type of jumps (see Section 4.2. by Berthonol et al., 2006). However for most of the methodologies used here, conditional volatilities $\sigma_t$ are functions of $\varepsilon_{t-1}$. With this modelling approach, the dynamics of volatility is a function of past jumps.

- A model mixing a conditionally Gaussian distribution with jumps:

$$r_t = \mu + \sigma_t \varepsilon_t + \sum_{i=0}^{N_t} x_i$$

with $N_t$ the number of jumps at time $t$, $x_i$, the scale of the $i$-th jump on date $t$ and $\varepsilon_t \sim N(0,1)$, $x_i \sim N(\mu_i, \sigma_i)$. With such an approach, the past volatility is no longer a function of past jumps, as $\varepsilon_t$ has been cleansed from the jump component. This latter component is assumed to be captured by $\sum_{i=0}^{N_t} x_i$, that is through a separated component. Tails events are captured by $\sum_{i=0}^{N_t} x_i$ while more normal days of trading are modelled through $\mu + \sigma_t \varepsilon_t$. In a continuous time setting, $\mu + \sigma_t \varepsilon_t$ is referred to as the continuous part of the process, while the jump part is the discontinuous part.

For each of these models, the structure for the continuous volatility is an Heterogenous Autoregressive Model, as presented by Corsi (2009). We propose to use this specification jointly with different high frequency measures of volatility:

- The realised variance for day $d$ that is given by the sum of squared intraday returns:

$$RV_{d,M} = \sum_{j=1}^{M} r_{d,j}^2$$

where the $r_{d,j}$ are intraday returns computed as $r_{d,j} = p_{d,j} - p_{d,j-1}$ for $j = 1, \ldots, M$. $p_{d,j}$ are intraday observations allowing to compute $M$ continuously compounded intraday returns each day.

- Barndorff-Nielsen and Shephard (2004)’s BPV measure, which is computed as the scaled summation of the product of adjacent absolute returns. Formally, BPV is defined as follows:

$$BPV_{d,m} = \varepsilon p \sum_{j=1}^{M} \left| r_{d,j+1} \right|$$

where $\varepsilon = 2^{p/2} \Gamma(\frac{1}{2}(2p+1)) = E(Z_{\mu})$ denotes the mean of the absolute value of standard normally distributed random variable $Z$. The BPV is a consistent estimator of integrated volatility and allows to decompose the realised volatility into its diffusive and non-diffusive parts. As the sampling frequency increases, the presence
of jumps should have no impact because the return representing the jump is multiplied by a non-jump return which tends to zero asymptotically. This is true in case of rare jumps (one each day) when the probability of two consecutive jumps is negligible.

- Nevertheless, the BPV can be upward or downward biased in empirical applications as the sampling frequency is not high enough to eliminate the influence of jumps (or in presence of zero-returns). This has motivated the need for alternative estimators which do not suffer from this weakness. Recently, Andersen et al. (2012) suggested the following Median Realised Volatility (MedRV) estimator:

\[ MedRV_N = \frac{\pi}{6-4\sqrt{3}} \left( \frac{N}{N-2} \right) \sum_{i=2}^{N-1} \text{med}(|\Delta Y_{i-1}|, |\Delta Y_i|, |\Delta Y_{i+1}|)^2 \]  

(6)

The MedRV estimator has two main advantages: first, the impact of jumps completely vanishes except in the case of two consecutive jumps (which is quite rare at the sampling frequencies used in our empirical application) and second the estimator is more robust to occurrence of zero-returns.\(^5\)

In Section 3, we discuss different empirical key elements related to the use of such volatility measures, such as the intraday sampling frequency. Using both the BPV and the MedRV, we present in the meantime a way to estimate jumps from intraday time series. For the time being, we assume that jumps are observable.

Now, we discuss the joint dynamics of volatility and returns. Maheu and McCurdy (2011) relate the conditional variance of daily returns \( \sigma_t^2 \) to the realised volatility estimator through a cross-equation restriction. Barndorff-Nielsen and Shephard (2002) and Andersen et al. (2003) show that under empirically realistic assumptions, the conditional variance of daily returns should equal the conditional expectation of quadratic variation or:

\[ \sigma_t^2 = E_{t-1}(QV_t) \]

(7)

where \( E_{t-1} \) stands for the conditional expectation at time \( t - 1 \), \( QV_t \) the quadratic variation, \( Var_{t-1}(r_t) \) and \( \sigma_t^2 \) the conditional variance of returns. Assuming that \( RV_t \) is an unbiased estimator of \( QV_t \), it follows that:

\[ \sigma_t^2 = E_{t-1}(RV_t). \]  

(8)

In other words, the one-period-ahead conditional expectation of the realised volatility should equal the ‘true’ conditional volatility assuming the unbiasedness of the realised volatility estimator. Under the assumption of a log-normal distribution for the realised volatility,\(^6\) the conditional expectation can then be expressed as:

\[ \sigma_t^2 = E_{t-1}(RM_t) = \exp \left( E_{t-1} \log(RM_t) + \frac{1}{2} Var_{t-1}(\log(RM_t)) \right) \]  

(9)

with \( RM_t \) a given realised measure (\( RV_t, BPV_t, MedRV_t \)).

We now turn to the specification of a predictive model for realised volatility. The HAR-RV model initially developed by Corsi (2009) has been used with success in a number of recent contributions (Andersen et al., 2007a; Andersen et al., 2007b; Corsi...
et al., 2008; Liu and Maheu, 2009, among others). The economic intuition behind this model is that different groups of investors have different investment horizons and consequently behave differently (see Muller et al., 1997) for the presentation of the HARCH original model relying on the Heterogeneous Hypothesis). The genuine HAR-RV model is formally a constrained AR(22) model using RV as the realised measures of variance but the HAR can naturally accommodate all realised measures and transformations of these measures. The HAR-RV model using daily, weekly and monthly realised volatility components may be written as follows:

$$\sqrt{RV_t} = \alpha_0 + \alpha_d \sqrt{RV_{t-1}} + \alpha_u (\sqrt{RV})_{-5t-4} + \alpha_n (\sqrt{RV})_{-22t-1} + \eta_t.$$  \hspace{1cm} (10)

The error term $\eta_t$ is chosen to fit the distribution of the residuals, but could very well be modelled as a GARCH error (see Corsi et al., 2008, Bollerslev et al., 2009). Since the logarithmic transformation exhibits superior forecasting performance (Andersen et al., 2007a; Andersen et al., 2007b), we retain the following specification:

$$\log(RV_t) = \omega + \phi_1 \log(RV_{t-1}) + \phi_2 \log(RV_{t-5})$$  \hspace{1cm} (11)

$$+ \phi_3 \log(RV_{t-22}) + \gamma \epsilon_{t-1} + \eta \nu_t \sim NID(0,1)$$  \hspace{1cm} (12)

The error term in the volatility equation is assumed to follow a standard Gaussian, as it is well-known since Andersen et al. (2001a, 2001b) that the logarithmic transformation of the realised volatility is normally distributed. The latter specification captures asymmetries coming from two distinct sources: leverage effects (through $\gamma$) and unconditional asymmetry (with the mixture of normals).

Densities forecasts using intraday futures will be compared with forecasts obtained with the traditional EGARCH model based on daily data:

$$r_t = \mu + \sigma_t \epsilon_t, \nu_t \sim NID(0,1)$$  \hspace{1cm} (13)

$$\log(\sigma_t^2) = \omega + \beta \log(\sigma_{t-1}^2) + \gamma \epsilon_{t-1} + \xi | \epsilon_{t-1} |$$  \hspace{1cm} (14)

To make comparisons easier, a leverage term is introduced in the volatility equation. Depending on the standardised return innovation $\epsilon_t$ in the return equation, the volatility is impacted through the coefficient $\gamma$. The estimated value of this coefficient is generally found to be negative. Indeed, an unexpected fall in returns translates into a positive impact on the level of volatility.

### 2.2 Parameters estimation of the time series models

The models presented in Section 2.1 are similar to those presented by Maheu and McCurdy (2011). As in their case, the estimation is performed by maximum likelihood: the estimation is possible as both the returns and the volatility are observed. On top of that – and in a similar fashion to Maheu and McCurdy (2011) – we assume that $\epsilon_t$ and $\nu_t$, that is the disturbances respectively associated to the returns and volatility, are uncorrelated. As it is well-known since Bertholon et al. (2006), the mixture of normals yields to estimates by QML which are very close to the true distribution.
We discuss rapidly the estimation of the model presented at equations (3)–(12,13).

Let \( \bar{r}_t \) be the ex-jump return, that is

\[
\bar{r}_t = r_t - \sum_{i=0}^{N_t} x_{i,t}.
\]

Let \( \Omega_t \) be vector containing the following three processes:

\[
\Omega_t = (\bar{r}_t, \sigma_t, J_t),
\]

where \( J_t \) is the jump component, that is \( \sum_{i=0}^{N_t} x_{i,t} \). The estimation of the parameters driving the joint behavior of the three processes can be obtained by maximising their joint likelihood. The conditional joint density given the past observation of \( \Omega_t \) can be written as follows:

\[
f(\bar{r}_t, \sigma_t, J_t | \Omega_{t-1}) = f(\bar{r}_t, \sigma_t | \Omega_{t-1}, \sigma_t)g(\sigma_t, J_t | \Omega_{t-1})
\]

\[
= f(\bar{r}_t | \Omega_{t-1}, \sigma_t)g(\sigma_t, J_t | \Omega_{t-1})h(J_t | \Omega_{t-1}, \sigma_t).
\]

In equations (20) and (21), \( f(\cdot), g(\cdot) \) and \( h(\cdot) \) are marginal densities. Equation (21) is obtained as \( \bar{r}_t \) and \( J_t \) are assumed to be indendent. The maximisation of the joint loglikelihood then clearly amounts to maximising each of its three components independently, given that they do not share common parameters.

Maximum likelihood estimates are thus obtained through the following expressions:

\[
\max_{\theta} \frac{1}{T} \sum_{t=1}^{T} \log f(\bar{r}_t | \Omega_{t-1}, \sigma_t)
\]

\[
\max_{\theta} \frac{1}{T} \sum_{t=1}^{T} \log g(\sigma_t, J_t | \Omega_{t-1})
\]

\[
\max_{\theta} \frac{1}{T} \sum_{t=1}^{T} \log h(J_t | \Omega_{t-1}, \sigma_t).
\]

As jumps are i.i.d. random variables, it is possible to obtain closed form expressions for the model’s parameters estimates:

\[
\hat{\lambda} = \frac{1}{T} \sum_{t=1}^{T} N_t
\]

\[
\hat{\mu}_t = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=0}^{N_t} x_{i,t}
\]

\[
\hat{\sigma}^2_t = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=0}^{N_t} (x_{i,t} - \hat{\mu}_t)^2.
\]
Such a straightforward estimation approach is made possible by two key elements: first, the fact that the volatility $\sigma_t$ is included within the filtration at time $t$; second, we assume that $\sum_{i=t}^{h} N_i$ is observable and unrelated to the other components in the dynamics of $r_t$.

With such an approach, we obtain a realistic split between the contributions of jumps and those of volatility to the evolutions of $r_t$.

### 2.3 Density forecasting power comparison

To compare the various models presented earlier, we rely on out-of-sample density forecasting exercises. This subsection presents the empirical approach that we retained.

To compare density forecasts between the standard EGARCH model and the bivariate model of daily returns and HAR, we use a criteria known as the predictive likelihood (or logarithmic score). The average predictive likelihood over the out-of-sample observations $t = \tau + k_{\text{max}}, \ldots, T - k$ is:

$$D_{\mathcal{M},k} = \frac{1}{T - \tau - k_{\text{max}} + 1} \sum_{\tau+1}^{T-k+1} \log f_{\mathcal{M},k} (r_{\tau+k} | \Phi_t, \theta), \quad k \geq 1$$

with $f_{\mathcal{M},k} (x | \Phi_t, \theta)$ the $k$-period ahead predictive density for model $\mathcal{M}$, given $\Phi_t$ and parameter $\theta$, evaluated at the realised return $x = r_{\tau+k}$. Therefore, better forecasts will translate into larger $D_{\mathcal{M},k}$.

We use a rolling window scheme to evaluate the predictive power of our forecasts. As by Maheu and MacCurdy (2011), compute the 1–60 days ahead forecasts for each window. We thus obtain $T - \tau - k_{\text{max}} + 1$ data blocks for each asset. For each block, we compute the predictive likelihoods and then average over all blocks.

To evaluate the relative accuracy of competing forecasts, we rely on the test statistics developed by Diebold and Mariano (1995) in the context of the comparison of density forecasts (Amisano and Giacomini, 2007). The null hypothesis is that predictive likelihood forecasts of horizon $h$ and models $\mathcal{A}$ and $\mathcal{B}$ have the same performance. The test statistics is given by:

$$t'_{\mathcal{A},\mathcal{B}} = \frac{(D_{\mathcal{A},k} - D_{\mathcal{B},k})}{(\sigma_{\mathcal{A},k} - \sigma_{\mathcal{B},k})} \sqrt{T - \tau - k_{\text{max}} + 1}$$

which is asymptotically standard normal. As for the interpretation of the Diebold and Mariano (1995) (DM) test, a significant positive (negative) estimated value rejects the null of equal performance between competing forecasts and provides evidence in favour of model $\mathcal{A}$ ($\mathcal{B}$).

An additional issue arises when computing the Amisano and Giacomini (2007)’s test statistics when the jumps are explicitly incorporated within the data generating process of $r_t$. Indeed, as future jumps are unobservable, we need to compute the conditional distribution of $r_{\tau+h}$ conditionally upon the information available at time $t$.

$$f(r_{\tau+h} | \mathcal{F}_t) = \sum_{i=0}^{h} f(r_{\tau+h} | \mathcal{F}_t, N_{\tau+h} = i) \times P(N_{\tau+h} = i),$$

(27)
where \( P(N_{t+h} = i) \) is the probability that the number of jumps between the date \( t + h - 1 \) and \( t + h \) is equal to \( i \), given the estimated parameters. Given the estimated values for the average number of jumps by working day, we approximate the previous quantity by truncating the previous infinite sum. We compute it from \( i = 0 \) to 10. Beyond this threshold, the results of the numerical tests remain qualitatively unaffected. In the various HAR-based models considered here, \( f(r_{t+h} | \mathcal{F}_t, N_{t+h} = i) \) is computed as by Maheu and McCurdy (2011) and in the previously mentioned cases that are not based on an explicit modelling of jumps. We also follow Maheu and McCurdy (2011) for the simulations, computations and the use of the Newey-West long-run variance (HAC) again to make our results comparable with theirs.

3 Empirical results

In this section, we present first the data used. Then, we detail the procedures used to find the optimal sampling frequency when using intraday data and to extract jumps. Finally, we comment in details the results obtained for all types of models (with/without jumps), along with the comparison of density forecasts.

3.1 Dataset

We purchased high-frequency futures data from TickData.\(^{12}\) While energy futures are very liquid (and are therefore suitable for using realised estimators), we need to remove days where the trading activity has not been sufficient to compute these estimators. To this end, we filter our time series with respect to three parameters: the length of the trading period in the day, the number of zero-returns and the number of transactions.

The West Texas Intermediate (WTI) light sweet crude oil futures contract is traded on the New York Mercantile Exchange (NYMEX) now a branch of CME. The period considered is from 8 October 2001 to 15 January 2010. The WTI contract is one of the most traded futures contract in the world. Using the procedure of rolling over front month futures contracts\(^{13}\), the total number of ticks for the continuous time series of the front month contract is equal to 52,099,419. The trading period for the WTI futures is from 9:00 AM to 2:30 PM, which should provide 60 intraday returns each day (54 intraday returns for the September 2001/January 2007 period where trading began at 10:00 AM). We remove days with less than 60 (54) intraday returns, days with more than 15 zero-returns and days with less than 700 registered ticks. The number of observations is therefore reduced from 2140 days to 2081 days when all these requirements are met. The mean number of trades is equal to 25,035. This figure is very different before and after mid-2006, which is mainly due to the launch of electronic trading.\(^{14}\)

Figure 1 displays the time-series of WTI futures, along with the open-to-close log-returns.\(^{15}\)

We also present summary statistics in Table 1. We observe that the realised volatility and the bi-power variation measures present non-zero skewness and excess kurtosis.\(^{16}\) These descriptive statistics therefore reveal a ‘fat tailed’ distribution. The logarithm transformations of these quantities are nearly Gaussian, which is a common finding since Andersen et al. (2001a, 2001b) among others.\(^{17}\)
Figure 1  Raw time-series (top panel) and open-to-close standardised log-returns (bottom panel) (see online version for colours)

Table 1  Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>$W_T$</th>
<th>$R_T$</th>
<th>$R_{VT}$</th>
<th>$BPV_T$</th>
<th>$MedRV_T$</th>
<th>$log(R_{VT})$</th>
<th>$log(BPV_T)$</th>
<th>$log(MedRV_T)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1445</td>
<td>0.3061</td>
<td>0.2954</td>
<td>0.2896</td>
<td>–8.1738</td>
<td>–2.2469</td>
<td>–2.860</td>
<td></td>
</tr>
<tr>
<td>Standard dev.</td>
<td>0.3846</td>
<td>0.0061</td>
<td>0.0057</td>
<td>0.0055</td>
<td>0.6940</td>
<td>0.6971</td>
<td>0.6975</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>–1.724</td>
<td>3.8354</td>
<td>3.9463</td>
<td>3.8867</td>
<td>0.6875</td>
<td>0.6670</td>
<td>0.6506</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.2218</td>
<td>19.9500</td>
<td>22.0793</td>
<td>20.8598</td>
<td>0.7653</td>
<td>0.7854</td>
<td>0.7764</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2058</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

Notes:  Mean values are given as the mean of the annualised squared root values.
3.2 Optimal sampling frequency and jump detection

This subsection presents empirical technicalities regarding the intraday volatility and jumps measures: (a) the determination of the optimal sampling frequency for each asset; (b) the jump detection and (c) the sequential jump detection procedures.

3.2.1 Optimal sampling frequency

For these three estimators of realised volatility, theory suggests that returns should be computed at the highest possible frequency, so that estimators converge asymptotically towards the true conditional volatility. However, it is well-known since Andersen and Bollerslev (1997, 1998) and Taylor and Xu (1997) that microstructure noise (due to price discreteness, bid-ask spread, non-synchronous trading, etc.) may impact the realised volatility estimator at high frequency.

To deal with this issue while making our results comparable with the rest of the literature, we follow the five minutes ‘rule-of-thumb’. As the WTI is a highly liquid asset, this sampling interval is adequate to make our realised measures not to be impacted by the noise.

We examine further this question for the WTI crude oil futures price series, which did not benefit from such an analysis in previous research. In Figures 2 and 3, we report the volatility signature plot for oil futures. This analysis is crucial as the trading activity dramatically increased in September 2006 following the generalisation of electronic trading (see Section 3.1) and may result in different noise structure before and after this event. These graphs confirm that the standard five minute sampling frequency seems to be appropriate in this case as well. Additional details on the tests conducted on the WTI futures contract are given in the Appendix A.

Figure 2 Volatility signature plot for the oil futures contract using front month rollover and the realised volatility, bipower variation and median realised estimators (2001–2006) (see online version for colours)
3.2.2 Jump detection

Once the optimal sampling frequency is determined, realised volatility, $BPV$ and median realised volatility estimators are computed. The difference between $RV$ and a jump-robust estimator such as $BPV$ or $MedRV$ provides, when it is statistically significant, an estimate of the sum of squared jumps $\sum_{j=1}^{n(t)} \kappa^2(t_j)$ which have occurred during the period under investigation. Note that a small estimated value for a jump may not be actually a jump but a variation due to the continuous path of the stochastic process and the presence of a jump has thus to be formally tested. BNS (2004, 2006) develop such a testing framework using asymptotic theory on realised variance and multipower variations.\(^{19}\)

As Andersen et al. (2007a, 2007b) put it, ‘significant’ jumps may be identified by comparing realisations of test statistics to a standard normal distribution. They use the test statistic by Huang and Tauchen (2005) to examine the significance of a jump when the chosen jump-robust estimator is the $BPV$:

$$
Z_{J_{BPV}}(N,d) = \sqrt{N} \frac{(RV_{d,N} - BPV_{d,N})RV_{d,N}^{-1}}{\left(\left(\frac{2}{\bar{\sigma}_1^2} + \frac{5}{\bar{\sigma}_2^2} - 5\right) \max\{1, TQ_{d,N}, BPV_{d,N}^{-2}\}\right)^{1/2}}
$$

with $TQ$ the realised tripower quarticity, which converges in probability to the integrated quarticity. The ratio-statistic in equation (28) has reasonable power against several empirically realistic calibrated stochastic volatility jump diffusion models (Andersen et al., 2007a; Andersen et al., 2007b).

The test may be adapted to the $MedRV$ estimator as follows:

$$
Z_{J_{MedRV}}(N,d) = \sqrt{N} \frac{(RV_{d,N} - MedRV_{d,N})RV_{d,N}^{-1}}{\left(0.96 \max\{1, MedRQ_{d,N}, MedRV_{d,N}^{-2}\}\right)^{1/2}}
$$

(29)
with MedRQ an estimate of the integrated quarticity obtained by using the same methodology as for MedRV. Theodossiou and Zikes (2009) show by means of many simulations and empirical analysis that this test has better properties in the presence of jumps of finite or even infinite activity and zero-returns. If, as we will demonstrate, disentangling jumps from the continuous component help in forecasting the density of returns, we are particularly interested in investigating whether the better properties of the test base on MedRV compared to BPV will translate in an improvement in density forecasting as well. A preliminary analysis shows that for both samples ‘jumpy days’ are similar using MedRV instead of BPV, but that the magnitude of jumps is slightly different.

Figure 4 shows the realised volatility, the BPV and the jump component from BPV, as well as the MedRV and the jump component from MedRV for the WTI series ‘jumpy days’ are quite similar using MedRV instead of BPV, but that the magnitude of jumps is slightly different.

Figure 4  Realised volatility, bi-power variation with jump component, and median realised volatility with jump component for WTI futures (from top to bottom and left to right) (see online version for colours)
More interestingly, Table 2 provides statistics about the contribution of jumps to the total realised volatility for the WTI series and for different level of significance of the test.\textsuperscript{22} Our results for WTI are new but very similar to Huang and Tauchen (2005) and Andersen et al. (2007a, 2007b) for equities. Overall, these results point to the fact that jumps contribute to a significant part of the total return variation. Because it is well-known that jumps are not persistent while the continuous component is, considering jumps independently is likely to be rewarded in a forecasting exercise.

Table 2

<table>
<thead>
<tr>
<th>Contribution of jumps</th>
<th>WTI crude oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1% threshold</td>
<td>2.61%</td>
</tr>
<tr>
<td>0.5% threshold</td>
<td>3.37%</td>
</tr>
<tr>
<td>1% threshold</td>
<td>3.81%</td>
</tr>
<tr>
<td>5% threshold</td>
<td>5.06%</td>
</tr>
<tr>
<td>no threshold</td>
<td>6.88%</td>
</tr>
</tbody>
</table>

Table 3 gives the Kolmogorov-Smirnov statistics for different distributions. Results indicate that the Gaussian distribution does not provide a good fit to the series of returns standardised by realised volatility, BPV and median realised volatility. Given these results (where we do not specifically model the jumps), we assume that the conditional distribution of returns is a mixture of Gaussian distributions, in a manner closely related to Maheu and McCurdy (2011).

Table 3

<table>
<thead>
<tr>
<th>WTI</th>
<th>RV(_i)</th>
<th>BPV(_i)</th>
<th>MedRV(_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Student-t</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Generalised Hyperbolic</td>
<td>0.94</td>
<td>0.55</td>
<td>0.95</td>
</tr>
<tr>
<td>Normal Inverse Gaussian</td>
<td>0.69</td>
<td>0.75</td>
<td>0.95</td>
</tr>
<tr>
<td>Hyperbolic</td>
<td>0.26</td>
<td>0.95</td>
<td>0.44</td>
</tr>
<tr>
<td>Mixture of Gaussian</td>
<td>0.84</td>
<td>0.95</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Notes: The distributions considered are the Gaussian, the symmetric Student-t, the Generalised Hyperbolic (GH), the Normal Inverse Gaussian (NIG), the Hyperbolic (H) and the Mixture of Normals (MN). The values in the table are p-values.

3.2.3 Sequential jump detection

For the purpose of modelling the jump component independently from the continuous component, we need a precise measure of jumps including their size along with their sign. We obtain such a measure using the sequential jump detection procedure developed in Andersen et al. (2011a).\textsuperscript{23} The idea behind this procedure is simple. We first use the standard test presented below using either BPV or MedRV on intraday returns for day \(d\). If we do not reject the existence of a jump for day \(d\), we consider that the largest return in absolute value is a jump and remove it from intraday returns. We replace this ‘jump’ return with the mean of remaining intraday returns and run the jump test again to check the presence of another jump in the same day.
One may wonder whether, in light of the rarity of jumps, sequential detection is useful? To answer this question we first plot in Figure 5 transactions for the WTI futures contract. The left panel plots a day with exactly one jump, whereas the right panel plots a day with two or more jumps identified with the sequential procedure.

Figure 5  Jumpy days (see online version for colours)

As a second and more rigorous evidence, we provide in Table 4 statistics about the number of jumps we detect each day for the WTI series. These statistics support the view that days with more than a single jump are quite common.

Table 4  Descriptive statistics about jumps extracted using the sequential procedure by Andersen et al. (2011a)

<table>
<thead>
<tr>
<th>Jump descriptive statistics</th>
<th>WTI crude oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of jumps</td>
<td>329</td>
</tr>
<tr>
<td>Number of days with jump(s)</td>
<td>232</td>
</tr>
<tr>
<td>Percentage of days with jump(s)</td>
<td>11.15%</td>
</tr>
<tr>
<td>Average duration between two jumps (days)</td>
<td>6.33</td>
</tr>
<tr>
<td>Average duration between two days with jump(s)</td>
<td>8.97</td>
</tr>
<tr>
<td>Number of days with exactly 1 jump</td>
<td>102</td>
</tr>
<tr>
<td>Percentage of days with exactly 1 jump</td>
<td>6.86%</td>
</tr>
<tr>
<td>Number of days with exactly 2 jumps</td>
<td>48</td>
</tr>
<tr>
<td>Percentage of days with exactly 2 jumps</td>
<td>2.31%</td>
</tr>
<tr>
<td>Number of days with exactly 3 jumps</td>
<td>10</td>
</tr>
<tr>
<td>Percentage of days with exactly 3 jumps</td>
<td>0.48%</td>
</tr>
<tr>
<td>Number of days with exactly 4 jumps</td>
<td>7</td>
</tr>
<tr>
<td>Percentage of days with exactly 4 jumps</td>
<td>0.34%</td>
</tr>
<tr>
<td>Number of days with exactly 5 jumps</td>
<td>2</td>
</tr>
<tr>
<td>Percentage of days with exactly 5 jumps</td>
<td>0.09%</td>
</tr>
</tbody>
</table>

Notes:  The chosen significance threshold is 1%.
Using the sequential detection procedure with both the BPV and the MedRV procedure, we obtain for each asset a series of jumps along with their sign and size which will be used in the econometric estimation below. As a rapid check of our results, we compare squared jumps with the squared component resulting from the difference between RV and the jump-robust measure (when significant) and confirm the importance of the sequential procedure.

3.3 Estimation results

We estimate the model on a rolling window of 1,260 daily observations. Concerning the average volatility level, we obtain the expected result that the BPV and MedRV estimators are less volatile than the ‘naive’ RV estimator (Table 1). For the WTI futures, we find evidence of leptokurticity (but globally less for log(BPV) and log(MedRV)).

Concerning the Kolmogorov-Smirnov statistics for daily open-to-close standardised returns, we uncover the adequation to various distributions. The returns are clearly non-Gaussian when looking at the skewness and kurtosis statistics in Table 3. We unambiguously reject the adequation to the Gaussian law. Due to the asymmetric patterns present in the data, the Student-t distribution is rejected as well. Interestingly, we accept various distributions which account for asymmetry such as the Generalised Hyperbolic and the Mixture of Gaussian distributions. In the estimation of the bivariate models, we focus on the Mixture of Normals because its parameters are easily interpretable in economic terms. In addition, its estimation is easier numerically and thus more adapted to a rolling estimation scheme.

Concerning the presence of jumps, we find for all assets a contribution of jumps to the volatility level (Tables 2 and 4). The contribution of jumps is equal to 6.88% for the WTI futures contract with no threshold for the detection as by Huang and Tauchen (2005). In terms of frequency, the number of days with jumps is quite important and strictly positive for crude oil. The most common situation is when we observe the occurrence of one jump per day. Furthermore, we note that the WTI futures contract is characterised by a very high frequency of jumps (with the percentage of days with jumps superior to 11%). The WTI futures records 44% of days with more than one jump over the sample interval.

Moving to the estimates of the bivariate model in Table 5, we focus our comments on the most interesting parameters. For the WTI futures contract, the leverage effect coefficient $\gamma$ is closer to zero when relying on the MedRV or BPV estimators. When looking at the persistence of the HAR components, we find that the $\phi_1$, $\phi_2$ and $\phi_3$ coefficients vary depending on the distributions and the measures of volatility. In the case of the WTI futures contract, the $\phi_1$ and $\phi_3$ components lose in persistence when one moves from the RV estimator to the BPV and MedRV estimators. However, the $\phi_2$ component gains relatively in persistence. These effects can be seen as being very specific to the sample data. In addition, the mixture parameter $\phi$ also varies depending on the volatility measure used. With the WTI futures contract, we find that $\hat{\phi} = 0.449$ with the RV estimator and $\hat{\phi} = 0.486$ with the BPV estimator.
Table 5

Forecasting the density of returns in crude oil futures markets

<table>
<thead>
<tr>
<th>EGARCH</th>
<th>$\omega$</th>
<th>$\alpha$</th>
<th>$\theta$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>$-0.563$</td>
<td>$-0.078$</td>
<td>$0.104$</td>
<td>$0.938$</td>
</tr>
<tr>
<td>Standard dev.</td>
<td>$0.404$</td>
<td>$0.029$</td>
<td>$0.014$</td>
<td>$0.052$</td>
</tr>
<tr>
<td>Skewness</td>
<td>$-0.433$</td>
<td>$-0.723$</td>
<td>$-0.525$</td>
<td>$-0.458$</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>$-1.014$</td>
<td>$-0.292$</td>
<td>$0.728$</td>
<td>$-0.970$</td>
</tr>
<tr>
<td>5% quantile</td>
<td>$-1.240$</td>
<td>$-0.128$</td>
<td>$0.075$</td>
<td>$0.849$</td>
</tr>
<tr>
<td>95% quantile</td>
<td>$-0.149$</td>
<td>$-0.047$</td>
<td>$0.123$</td>
<td>$0.990$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HAR-RV-Gaussian</th>
<th>$\omega$</th>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\phi_3$</th>
<th>$\gamma$</th>
<th>$\mu$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>$-0.736$</td>
<td>$0.076$</td>
<td>$0.543$</td>
<td>$0.294$</td>
<td>$-0.041$</td>
<td>$0.001$</td>
<td>$0.422$</td>
</tr>
<tr>
<td>Standard dev.</td>
<td>$0.444$</td>
<td>$0.007$</td>
<td>$0.033$</td>
<td>$0.026$</td>
<td>$0.003$</td>
<td>$0.000$</td>
<td>$0.004$</td>
</tr>
<tr>
<td>Skewness</td>
<td>$-0.072$</td>
<td>$0.104$</td>
<td>$-0.299$</td>
<td>$-0.267$</td>
<td>$-0.225$</td>
<td>$0.260$</td>
<td>$-0.257$</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>$-1.842$</td>
<td>$0.484$</td>
<td>$-1.487$</td>
<td>$-0.954$</td>
<td>$0.138$</td>
<td>$-1.016$</td>
<td>$-0.898$</td>
</tr>
<tr>
<td>5% quantile</td>
<td>$-1.324$</td>
<td>$0.064$</td>
<td>$0.490$</td>
<td>$0.250$</td>
<td>$-0.047$</td>
<td>$0.000$</td>
<td>$0.414$</td>
</tr>
<tr>
<td>95% quantile</td>
<td>$-0.255$</td>
<td>$0.086$</td>
<td>$0.582$</td>
<td>$0.331$</td>
<td>$-0.036$</td>
<td>$0.001$</td>
<td>$0.428$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HAR-BPV-Gaussian</th>
<th>$\omega$</th>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\phi_3$</th>
<th>$\gamma$</th>
<th>$\mu$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>$-0.698$</td>
<td>$0.085$</td>
<td>$0.543$</td>
<td>$0.289$</td>
<td>$-0.038$</td>
<td>$0.001$</td>
<td>$0.410$</td>
</tr>
<tr>
<td>Standard dev.</td>
<td>$0.409$</td>
<td>$0.011$</td>
<td>$0.051$</td>
<td>$0.017$</td>
<td>$0.004$</td>
<td>$0.000$</td>
<td>$0.003$</td>
</tr>
<tr>
<td>Skewness</td>
<td>$-0.080$</td>
<td>$-0.121$</td>
<td>$-0.089$</td>
<td>$-0.883$</td>
<td>$-0.259$</td>
<td>$0.260$</td>
<td>$-0.262$</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>$-1.850$</td>
<td>$-0.099$</td>
<td>$-1.784$</td>
<td>$1.708$</td>
<td>$-0.204$</td>
<td>$-1.016$</td>
<td>$-0.106$</td>
</tr>
<tr>
<td>5% quantile</td>
<td>$-1.232$</td>
<td>$0.066$</td>
<td>$0.476$</td>
<td>$0.258$</td>
<td>$-0.045$</td>
<td>$0.000$</td>
<td>$0.405$</td>
</tr>
<tr>
<td>95% quantile</td>
<td>$-0.258$</td>
<td>$0.101$</td>
<td>$0.605$</td>
<td>$0.314$</td>
<td>$-0.031$</td>
<td>$0.001$</td>
<td>$0.414$</td>
</tr>
</tbody>
</table>
Table 5  
Estimation results for WTI Futures (continued)

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Average</th>
<th>Standard dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>5% quantile</th>
<th>95% quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAR-MedRV-Gaussian</td>
<td>$\omega$</td>
<td>-0.710</td>
<td>0.089</td>
<td>0.524</td>
<td>0.304</td>
<td>-0.036</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\phi_1$</td>
<td>0.419</td>
<td>0.020</td>
<td>0.061</td>
<td>0.018</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>$\phi_2$</td>
<td>-0.107</td>
<td>0.152</td>
<td>-0.123</td>
<td>-0.347</td>
<td>-0.124</td>
<td>0.260</td>
</tr>
<tr>
<td></td>
<td>$\phi_3$</td>
<td>-1.818</td>
<td>-1.498</td>
<td>-1.813</td>
<td>-0.031</td>
<td>-0.365</td>
<td>-1.016</td>
</tr>
<tr>
<td></td>
<td>$\gamma$</td>
<td>-1.277</td>
<td>0.061</td>
<td>0.442</td>
<td>0.275</td>
<td>-0.042</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>$\mu_1$</td>
<td>-0.262</td>
<td>0.117</td>
<td>0.595</td>
<td>0.331</td>
<td>-0.029</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\eta$</td>
<td>0.413</td>
<td>0.419</td>
<td>0.020</td>
<td>0.061</td>
<td>0.004</td>
<td>0.000</td>
</tr>
</tbody>
</table>

| HAR-RV-Mixture   | $\omega$ | -0.693  | 0.100         | 0.524    | 0.298    | -0.038      | 0.449        |
|                  | $\phi_1$  | 0.395   | 0.027         | 0.068    | 0.049    | 0.005       | 0.173        |
|                  | $\phi_2$  | -0.232  | -0.112        | -0.493   | -0.215   | -0.004      | -0.654       |
|                  | $\phi_3$  | -1.403  | 2.986         | 1.660    | 3.405    | 3.367       | -0.245       |
|                  | $\gamma$  | -1.282  | 0.060         | 0.413    | 0.223    | -0.046      | 0.098        |
|                  | $\mu_1$   | -0.201  | 0.140         | 0.616    | 0.373    | -0.030      | 0.677        |
|                  | $\mu_2$   | -0.763  | 1.057         | 0.626    | 1.238    | 0.001       | 0.422        |
|                  | $\sigma_1$| 0.486   | 0.411         | 0.054    | 0.009    | 0.219       | 0.324        |
|                  | $\sigma_2$| -0.319  | 0.041         | -0.474   | -0.407   | -0.182      | -0.636       |
|                  | $\sigma_3$| -0.898  | 2.343         | 1.962    | 2.900    | 8.250       | 2.359        |
|                  | $\mu$     | -0.144  | 0.190         | 0.656    | 0.344    | -0.020      | 0.769        |
|                  | $\eta$    | 0.417   | 0.410         | 0.119    | 0.540    | 0.263       | -0.032       |

| HAR-BPV-Mixture  | $\omega$ | -0.695  | 0.119         | 0.540    | 0.263    | -0.032      | 0.486        |
|                  | $\phi_1$  | 0.411   | 0.041         | 0.068    | 0.054    | 0.009       | 0.219        |
|                  | $\phi_2$  | -0.319  | 0.464         | -0.474   | -0.407   | -0.182      | -0.636       |
|                  | $\phi_3$  | -0.898  | 2.343         | 1.962    | 2.900    | 8.250       | 2.359        |
|                  | $\gamma$  | -1.348  | 0.060         | 0.424    | 0.175    | -0.044      | 0.046        |
|                  | $\mu_1$   | -0.144  | 0.190         | 0.656    | 0.344    | -0.020      | 0.769        |
|                  | $\mu_2$   | -0.763  | 1.057         | 0.626    | 1.238    | 0.001       | 0.422        |
|                  | $\sigma_1$| 0.486   | 0.411         | 0.054    | 0.009    | 0.219       | 0.324        |
|                  | $\sigma_2$| -0.319  | 0.041         | -0.474   | -0.407   | -0.182      | -0.636       |
|                  | $\sigma_3$| -0.898  | 2.343         | 1.962    | 2.900    | 8.250       | 2.359        |
|                  | $\mu$     | -0.144  | 0.190         | 0.656    | 0.344    | -0.020      | 0.769        |
|                  | $\eta$    | 0.417   | 0.410         | 0.119    | 0.540    | 0.263       | -0.032       |
Table 5: Estimation results for WTI Futures (continued)

<table>
<thead>
<tr>
<th>Model</th>
<th>$\omega$</th>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\phi_3$</th>
<th>$\gamma$</th>
<th>$\phi$</th>
<th>$\mu_1$</th>
<th>$\sigma_1$</th>
<th>$\mu_2$</th>
<th>$\sigma_2$</th>
<th>$\mu$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HAR-MedRV-Mixture</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>-0.666</td>
<td>0.110</td>
<td>0.543</td>
<td>0.273</td>
<td>-0.030</td>
<td>0.518</td>
<td>-0.748</td>
<td>1.054</td>
<td>0.852</td>
<td>1.300</td>
<td>0.001</td>
<td>0.413</td>
</tr>
<tr>
<td>Standard Dev.</td>
<td>0.583</td>
<td>0.036</td>
<td>0.062</td>
<td>0.041</td>
<td>0.007</td>
<td>0.164</td>
<td>0.239</td>
<td>0.100</td>
<td>0.299</td>
<td>0.274</td>
<td>0.000</td>
<td>0.003</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.261</td>
<td>0.261</td>
<td>-0.610</td>
<td>0.584</td>
<td>-0.709</td>
<td>-0.841</td>
<td>-1.572</td>
<td>5.561</td>
<td>-0.375</td>
<td>0.972</td>
<td>0.260</td>
<td>-0.039</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.989</td>
<td>2.461</td>
<td>3.723</td>
<td>1.194</td>
<td>6.061</td>
<td>0.689</td>
<td>4.999</td>
<td>31.912</td>
<td>0.094</td>
<td>0.898</td>
<td>-1.016</td>
<td>-0.854</td>
</tr>
<tr>
<td>5% quantile</td>
<td>-1.247</td>
<td>0.051</td>
<td>0.447</td>
<td>0.215</td>
<td>-0.040</td>
<td>0.186</td>
<td>-1.162</td>
<td>1.007</td>
<td>0.280</td>
<td>0.952</td>
<td>0.000</td>
<td>0.409</td>
</tr>
<tr>
<td>95% quantile</td>
<td>-0.161</td>
<td>0.167</td>
<td>0.645</td>
<td>0.346</td>
<td>-0.021</td>
<td>0.743</td>
<td>-0.427</td>
<td>1.066</td>
<td>1.274</td>
<td>1.820</td>
<td>0.001</td>
<td>0.417</td>
</tr>
<tr>
<td><strong>HAR-BPV-Jump</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>-0.704</td>
<td>0.084</td>
<td>0.544</td>
<td>0.288</td>
<td>-0.038</td>
<td>0.001</td>
<td>0.410</td>
<td>0.149</td>
<td>-0.001</td>
<td>0.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Dev.</td>
<td>0.412</td>
<td>0.011</td>
<td>0.051</td>
<td>0.018</td>
<td>0.004</td>
<td>0.000</td>
<td>0.003</td>
<td>0.014</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.106</td>
<td>-0.072</td>
<td>-0.015</td>
<td>-0.549</td>
<td>-0.256</td>
<td>0.260</td>
<td>-0.260</td>
<td>-0.308</td>
<td>-0.199</td>
<td>0.479</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-1.814</td>
<td>0.113</td>
<td>-1.676</td>
<td>0.548</td>
<td>-0.167</td>
<td>-1.016</td>
<td>-0.115</td>
<td>-0.088</td>
<td>-0.694</td>
<td>-1.117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5% quantile</td>
<td>-1.273</td>
<td>0.065</td>
<td>0.478</td>
<td>0.254</td>
<td>-0.044</td>
<td>0.000</td>
<td>0.405</td>
<td>0.121</td>
<td>-0.003</td>
<td>0.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% quantile</td>
<td>-0.261</td>
<td>0.101</td>
<td>0.612</td>
<td>0.315</td>
<td>-0.031</td>
<td>0.001</td>
<td>0.414</td>
<td>0.171</td>
<td>0.000</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HAR-MedRV-Jump</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>-0.719</td>
<td>0.088</td>
<td>0.523</td>
<td>0.303</td>
<td>-0.036</td>
<td>0.001</td>
<td>0.413</td>
<td>0.156</td>
<td>-0.002</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Dev.</td>
<td>0.419</td>
<td>0.020</td>
<td>0.060</td>
<td>0.018</td>
<td>0.004</td>
<td>0.000</td>
<td>0.003</td>
<td>0.015</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.112</td>
<td>0.129</td>
<td>-0.094</td>
<td>-0.143</td>
<td>-0.143</td>
<td>0.260</td>
<td>-0.013</td>
<td>-0.544</td>
<td>0.302</td>
<td>0.432</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-1.806</td>
<td>-1.473</td>
<td>-1.779</td>
<td>-0.264</td>
<td>-0.274</td>
<td>-1.016</td>
<td>-0.874</td>
<td>0.057</td>
<td>-0.827</td>
<td>-1.277</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5% quantile</td>
<td>-1.292</td>
<td>0.062</td>
<td>0.444</td>
<td>0.275</td>
<td>-0.042</td>
<td>0.000</td>
<td>0.409</td>
<td>0.123</td>
<td>-0.003</td>
<td>0.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% quantile</td>
<td>-0.265</td>
<td>0.116</td>
<td>0.596</td>
<td>0.333</td>
<td>-0.029</td>
<td>0.001</td>
<td>0.417</td>
<td>0.177</td>
<td>-0.001</td>
<td>0.012</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note that we can also compare the degree of activity of the jump component, i.e. it is possible to rank the assets depending on the intensity and the volatility level of jumps. In the case of the WTI futures contract, the parameter recording the intensity of jumps $\lambda$ is equal to 0.149. Finally, the WTI futures contract exhibits a very high volatility of jumps, with the coefficient $\sigma_x$ being equal to 0.010 for the HAR-BPV-Jump model. Thus, we uncover that there are different types of jumps (i.e. either frequent and small jumps or less frequent and large jumps) present in crude oil futures.

### 3.4 Forecast accuracy

In terms of forecast accuracy, we find that the bivariate models which take explicitly into account the jumps tend to perform better than the models based on a Gaussian distribution. In Table 6, the Diebold-Mariano test statistic is equal to 4.717 for the Realised Gaussian vs. Bipower Gaussian models at a 30 days horizon. Besides, we confirm the result by Maheu and McCurdy (2011) that the EGARCH model is dominated by the Realised Gaussian model (with a Diebold-Mariano test statistic equal to $-3.874$ at the ten days horizon). More generally, at the five and ten days horizons, all non-Gaussian models are equivalent, whatever the volatility measure. At the five days horizon, all non-Gaussian models either are equivalent or dominate Gaussian models according to the Diebold-Mariano test statistics in Table 6. For the horizons superior to ten days, the Realised MN model dominates the Bipower MN and MedRV MN models. In addition, we note that the BPV Jump model dominates the MedRV model at all horizons.

The same insights can be gathered by looking at the average predictive likelihood for the density of the WTI asset at various horizons (upto 60 days) in Figure 6. When looking at these graphs, recall that for the interpretation of the Diebold-Mariano test statistic, a significant positive (negative) estimated value rejects the null of equal performance between competing forecasts and provides evidence in favour of model $A$ ($B$). These graphs confirm the highly superior forecast accuracy of models based on intraday data for the WTI series. This was the main result by Maheu and McCurdy (2011) and we confirm their findings. BPV and MedRV estimators have a good performance for the WTI in comparison with models using realised volatility estimated using the mixture of normals. This indicates that the jump component either included in the total realised volatility (using the mixture of normals) or modelled separately provides information in forecasting the density of returns. The superiority of the BPV Jump and the MedRV Jump is evident at all horizons.

Departing from Maheu and McCurdy (2011), our results therefore tend to provide a deeper understanding of the effects at stake when decomposing the jump and continuous components of volatility. Namely, modelling explicitly jumps is of primary importance to achieve better performances with bivariate models, while decomposing between jumps and the continuous component volatility appears of secondary importance.
Table 6  Average Diebold–Mariano (1995) pairwise test statistic with five days, ten days, 30 days and 60 days horizon for WTI futures

Table 6  Average Diebold-Mariano (1995) pairwise test statistic with five days, ten days, 30 days and 60 days horizon for WTI futures (continued)

<table>
<thead>
<tr>
<th>60 day horizon</th>
<th>EGARCH</th>
<th>Realised Gaussian</th>
<th>Bipower Gaussian</th>
<th>MedRV Gaussian</th>
<th>Realised MN</th>
<th>Bipower MN</th>
<th>MedRV MN</th>
<th>Jump MN</th>
<th>Jump</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realised Gaussian</td>
<td>–</td>
<td>–</td>
<td>3.009</td>
<td>2.874</td>
<td>–1.087</td>
<td>–1.624</td>
<td>0.204</td>
<td>–2.416</td>
<td>–0.439</td>
</tr>
<tr>
<td>Bipower Gaussian</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>2.540</td>
<td>–1.588</td>
<td>–2.090</td>
<td>–0.767</td>
<td>–23.491</td>
<td>–3.515</td>
</tr>
<tr>
<td>MedRV Gaussian</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–1.730</td>
<td>–2.189</td>
<td>–1.049</td>
<td>–11.892</td>
<td>–6.148</td>
</tr>
<tr>
<td>Realised MN</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–0.529</td>
<td>4.604</td>
<td>0.174</td>
<td>0.510</td>
</tr>
<tr>
<td>Bipower MN</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4.075</td>
<td>0.327</td>
<td>0.688</td>
<td></td>
</tr>
<tr>
<td>MedRV MN</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–0.888</td>
<td>–0.327</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipower Jump</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–1.294</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MedRV Jump</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table presents the results of the average Diebold-Mariano (1995) pairwise test statistic with five days, ten days, 30 days and 60 days horizon. The table reads the following way: a significant positive (negative) estimated value rejects the null of equal performance between competing forecasts, and provides evidence in favor of model A (B).

Figure 6  Average predictive likelihood based on Amisano and Giacomini (2007) for the density of WTI futures (see online version for colours)
4 Conclusion

This paper has examined the forecasting power of jumps in addition, specifically, to the continuous component when the density of returns is the variable of interest. Whilst numerous authors have considered the informational content of continuous vs. jump components for volatility forecasting, none have thought to address the particular question of density forecasting. Detection of such information for different classes of assets would indicate the ability of new econometric models to anticipate the evolution of density returns in a fundamentally different way compared to more traditional forecasting models.

Our results unveil new effects regarding the importance of distinguishing between the continuous and jump components of volatility for the WTI Light Sweet Crude Oil Futures contract over 2001–2010. In this regard, this article specifically extends the findings by Maheu and McCurdy (2011) by considering various bivariate models with/without jumps. The main results may be summarised as follows. First, we confirm the findings by Maheu and McCurdy (2011) that intraday data yield better densities forecasts than daily data. Second and more importantly, we assess the importance of jumps when forecasting the density of returns by comparing jump-robust (BPV, median realised volatility) and non-robust measures of realised volatilities. Compared to the ‘naive’ measure of realised volatility, considering jumps specifically provides significant improvement on the accuracy of forecasts of return densities.

Two central concluding remarks arise. First, we have shown in this paper that the explicit modelling of jumps is central in the estimation of bivariate models in the fashion of Maheu and McCurdy (2011). Such an explicit modelling task yields better performances, as shown in our empirical application and with the forecast accuracy tests of all competing models. Second, discriminating between the continuous component of volatility and jumps appears comparatively less important across our estimates. From that perspectives, our results can be seen as an extension of the contribution by Maheu and McCurdy (2011).

Other jump detection techniques may be used (see Boudt et al., 2011; Christensen et al., 2010 among others). Boudt et al. (2011) in particular provide very interesting empirical results because their measure take the intraday periodicity into account and thus does not over-detect jumps in low-volatility periods and does not under-detect jumps in periods of high-volatility. Nevertheless, collectively taken, our results are sufficiently strong so that we can believe they would be robust to alternative jump detection methods.

Acknowledgements

Part of the paper was written while the author was visiting researcher at the Imperial College Business School (Imperial College London, UK). The author is thankful to Walter Distaso, Robert Kosowski, Diaa Noureldin, Dobrislav Dobrev and Filip Zikes for many insightful comments. Helpful comments were received from participants at the 2009 University Paris West Econometrics Workshop (France), the 18th International Conference on Forecasting Financial Markets (Marseille, France) and the 5th CSDA International Conference on Computational and Financial Econometrics (CFE’11, Birbeck College, University of London, UK).
References


Forecasting the density of returns in crude oil futures markets


Notes

1 Note we are not interested here in comparing the forecasting power of implied volatility vs. realised measures. This issue has been widely covered in the literature where future volatility is the variable of interest (Pong et al., 2004).

2 It has also advantages when residuals of univariate models are contemporaneously correlated as highlighted by Bollerslev et al. (2009). We do not consider here the possible correlation between errors in single equations, which is left for further research.

3 Note that the representation of the HAR-RV model by Corsi (2009) is linear. Running structural break tests appear to go beyond the scope of the current research, which is already heavy in terms of econometrics content. Structural breaks are left for further research and standalone papers on the topic of crude oil price forecasting.

4 This notation is used consistently in the paper.

5 We could use other estimators to obtain measures of integrated variance, such as QRV (Christensen et al., 2010) estimator, which are shown to be more robust in the presence of microstructure noise and zero-returns. A comparison of these estimators and their properties for density forecasting is beyond the scope of this paper and left for future research.

6 Empirical evidence of this hypothesis can be found in early contribution such as ABDL (2001a and b, 2003). Similar evidence for foreign exchange rates, futures markets, crude oil futures and the FTSE index may be found by Pong et al. (2004), Thomakos and Wang (2003), Wang et al. (2008) and Areal and Taylor (2002), respectively.

7 Forsberg and Ghysels (2007) and Ghysels and Sohn (2009), note that other power transformations may be used to model the dynamics of the realised volatility. These studies show that for a number of stochastic volatility processes used in the financial literature the absolute value of the realised volatility is a better predictor of the future realised volatility, particularly for longer horizons. We do not follow this approach here.
Forecasting the density of returns in crude oil futures markets

8 The optimal lag structure for the HAR model has been investigated by Craioveanu and Hillebrand (2010) who find that the genuine structure suggested by Corsi (2009) performs the best.

9 Note that the model does not allow the asymmetry to propagate into future volatility as in the EGARCH model.

10 As by Maheu and McCurdy (2011), we do not resort to Monte-Carlo simulations, which would be too computationally demanding in terms of numerical implementation, particularly in a rolling window setting that we adopt for the out-of-sample forecasting comparison.

11 The assumption of conditional independence does not lead to unconditional independence as the two equations are related through the leverage term. A more complete model allowing for conditional dependence has been developed by Bollerslev et al. (2009).

12 We refrain from purchasing additional years of data from TickData, which come at a hefty price. Besides, the estimation of our density forecasts models is very time-consuming and we do not want to impose a too high computational burden on our machine by adding several hundred thousands additional quotes. There is a trade-off to be made here between the realistically achievable objectives of the paper and the most sophisticated (highly parameterised) study on the topic.

13 The series is built such as each contract is rolled over 11 days before maturity. If this corresponds to a non-working day, we consider the day just before. This choice of rolling over futures contracts will not introduce significant bias in our estimates. Thus, we follow this approach because of its simplicity. We notice a slight maturity effect in the crude oil futures contract, which does not need to be modelled here. It may be considered more carefully for portfolio applications.

14 We have an average number of 2214 ticks per day during the 4 September 2001/31 August 2006 period vs. 57,054 ticks for the 1 September 2006/15 January 2010 period.

15 We choose to work with open-to-close returns because overnight returns have shown to follow a very different dynamics. In addition, including overnight returns may alter our analysis when standardising returns as we work with volatility computed with intraday transaction data.

16 Note for normally distributed random variable skewness is zero and kurtosis is three.

17 We come back on this issue when modelling the volatility using the log transformation. Goncalves and Meddahi (2011) suggest a new class of non-linear transformations based on the Box-Cox transformation which outperforms the log transformation in Monte Carlo simulations. We leave as an extension this possible transformation and follow the bulk of the empirical literature by considering the logarithm.

18 See Hansen and Lunde (2006) for a thorough discussion of this issue and Andersen et al. (2011b) for a theoretical and empirical analysis of the impact of microstructure noise on the forecast of realised volatility. To deal with this issue, we use staggered versions of BPV and MedRV as advocated by Huang and Tauchen (2005).

19 Veraart (2010) studies the limit theory of these estimators in the presence and absence of jumps.

20 Several other estimators for identifying jumps in the series, such as QRV (Christensen et al., 2010) estimator, which are shown to be more robust in the presence of microstructure noise and zero-returns, may be used. Theodossiou and Zikes (2009) provide a very complete treatment of existing jump detection tests as well as their relative performance in case of microstructure noise and/or jumps. In light of the good properties of the MedRV-based test, we focus on this alternative estimator.

21 Not reproduced here to conserve space, but available upon request to the authors.

22 The ‘no threshold’ case corresponds to the case measure of jump contribution by Huang and Tauchen (2005) where the difference between RV and BPV is taken directly.

23 A quite similar procedure is developed earlier by Andersen et al. (2007b).
A different result is obtained using mid-quotes data, which is, by nature, less prone to bid-ask bounds effects. Some researchers use quotes data which may support their choice of five-minute returns. We do not have access to this kind of data.

A similar pattern can be observed for BPV and is available upon request to the authors.

Recent contributions by Zhang et al. (2005) and Bandi and Russell (2008), among others, have proposed other methodologies to deal with noise, while ensuring that a maximum of data is used. Barndorff-Nielsen and Shephard (2008) provide a discussion and a unifying framework for these methodologies using kernel-based representation. These ideas have not been extended yet to the estimation of a jump-robust statistic (such as BPV) and cannot be applied here.
Appendix A: Optimal sampling frequency for WTI crude oil futures

Previous studies (Wang et al., 2008) relying on intraday data for the NYMEX crude oil futures contracts did not consider the issue of optimal sampling when estimating the realised volatility. These studies use the five minute returns interval popularised by Andersen and Bollerslev (1998). None of them attempts at looking at the relevance of this sampling frequency. If the choice of five minute returns has been validated in the empirical finance literature for S&P 500 futures, foreign exchange rates and some individuals stocks, it may not be adapted for the crude oil futures series.

Here, we show empirically that the five minute returns sampling interval is not relevant in the case of crude oil. We propose a rigorous methodology to determine the optimal sampling frequency using a rolling version of the test recently developed by Awartani et al. (2009). In light of this analysis, ten minute returns intervals for computing returns seem to provide the best balance between sufficiently fine sampling and a limited impact of the microstructure noise.

In Figure 7, we plot the average realised variance for the full sample of crude oil futures data (1 September 2001 to 28 February 2009) against the sampling frequency (in minutes). This graph has been coined as a ‘volatility signature plot’ in the empirical finance literature (see Andersen et al., 2001a; Andersen et al., 2001b; Andersen et al., 2003) among others). It allows to visualise the impact of increasing the sampling frequency. In view of Figure 7, it can be observed that the average realised variance is increasing with the sampling frequency.

This effect may be attributed to the impact of microstructure noise, whose volatility becomes predominant over the volatility of the asset price when the interval between
observations is diminishing. The empirical finance literature often relies on the informal inspection of the volatility signature plot to choose the optimal sampling frequency. Between sampling data at a higher frequency and limiting the impact of microstructure noise, we use here a more rigorous approach by applying the recent test proposed by Awartani et al. (2009) in a rolling version as advocated by the authors.

Awartani et al. (2009) implement two different statistics to test whether the market microstructure noise has a significant impact on a given realised volatility measure. The idea is to compare two different frequencies against each other and to investigate the significance of the impact of the noise at the highest frequency. Applied on a rolling basis, this procedure may be viewed as a formal and statistically robust version of the volatility signature plot. The first ZT-statistic is based on the standard estimator of variance (realised variance). The second ZT-statistic relies on the concept of BPV and as such is robust to the presence of jumps. At this point, we have no information about the behavior of the crude oil futures series in terms of jumps. Thus, we implement the second statistic, which may be written as follows:

$$ZT_{BPV}(M, N, \tau, d) = \sqrt{NT} \left( \frac{BPV_{\tau, M} - BPV_{\tau, N}}{\sqrt{\alpha QV_{\tau, N}}} \right)$$

with $\tau$ a fixed time span, $M$ and $N$ different sampling frequencies such that $\frac{N}{M} \to 0$ when $M, N \to \infty$. $QV_{\tau, N}$ is the scaled realised quadpower variation and $\alpha$ determined as in Barndorff-Nielsen and Shephard (2006). Awartani et al. (2009) run their simulations on five days blocks ($\bar{\tau} = \text{five days}$), which seems to provide enough observations to harness the power of the test. We follow thoroughly their approach.

**Figure 8** Statistical volatility signature plot for the oil futures contract using front month rollover and the bipower variation estimator according to the ZT-test statistics in Awartani et al. (2009).
In Figure 8, we plot the percentage of rejection of the null hypothesis ($H_0$: the noise has no impact) against the sampling frequency in minutes. We label this graph a ‘statistical volatility signature plot’. It may be viewed as the statistical equivalent of standard volatility signature plots. The optimal sampling frequency for the crude oil series can be detected at the 5% rejection level. To simplify calculation while ensuring a sufficient number of intraday observations, we choose a sampling interval of five minute returns to compute realised volatility and $BPV$ for crude oil futures. While it could be considered as conservative, this choice ensures a very limited and insignificant impact of microstructure noise.

The implementation of Awartani et al.’s (2009) test thus guarantees that the choice of five minute returns as the optimal sampling frequency for crude oil futures will not impact the accuracy of the forecasts obtained in Section 3.
The effect of oil prices on stock prices: a structural time series approach

Imad Moosa

School of Economics, Finance and Marketing,
RMIT University,
Melbourne, Victoria 3000, Australia
Email: imad.moosa@rmit.edu.au

Abstract: A structural time series model is estimated to examine the effect of oil prices on stock prices in three oil importing and four oil exporting countries. The results show that some missing variables affect the secular behaviour of, and that oil prices can explain cyclical variation in, stock prices. In all cases the effect of oil prices on stock prices turns out to be significantly positive. This finding is explained in terms of the positive effect running from economic activity to both oil prices and stock prices.

Keywords: oil prices; structural time series modelling; demand shocks; supply shocks; asymmetric responses; missing variables; TVP estimation; Kalman filter.


Biographical notes: Imad Moosa is currently a Professor of Finance at RMIT, Melbourne. Before taking on the present position, he was a Professor of Finance at Monash University and La Trobe University, and a Lecturer in Economics and Finance at the University of Sheffield. Prior to becoming an academic, he was a professional economist and a financial journalist for over ten years, and he also worked as an economist at the Financial Institutions Division of the Bureau of Statistics, International Monetary Fund. He has served in a number of advisory positions, including his role as an advisor to the US Treasury.

1 Introduction

It is intuitive to envisage that the effect of oil prices on aggregate stock prices (represented by a market index) depends on whether the underlying country is a net oil importer or a net oil exporter. For a net oil importer, higher oil prices mean higher costs of production, which should depress profitability and lead to lower stock prices. For a net oil exporter, higher oil prices boost government revenue, which should have a positive effect on stock prices, since government expenditure is the propelling force of economic activity in oil exporting countries (at least the countries examined in this paper). However, the available empirical evidence is mixed – for one reason because the effect depends on whether oil prices are driven by demand shocks or supply shocks as argued by Killian and Park (2009).
The consensus view on the effect of oil prices on stock prices in developed countries (invariably oil importing countries) is that rising oil prices are generally bearish for the stock market and vice versa. On 21 August 2006, the Financial Times attributed the decline of the US stock market to rising oil prices caused by concerns about political instability in the Middle East. On 12 October 2006, the same newspaper expressed the view that strong rallies in stock markets were due to a slide in crude oil prices that same day. This view follows from the fact that higher oil prices result in reduced profit margins for companies and depressed demand for consumer goods. However, higher oil prices also lead to the emergence of inflationary bursts, which means that if inflation and stock returns are positively related (at least in the long run, as the empirical evidence indicates), there should be a positive association between stock prices and oil prices. Furthermore, higher oil prices lead to higher profits for oil and oil-related companies, which should have a positive effect on the stock prices of these companies.

The conventional view that oil prices and stock prices are negatively correlated is not necessarily consistent with the stylised facts. For example, Walayat (2011) shows a graph of the Dow Jones index and crude oil price over a period going back to 2005, exhibiting mostly positive correlation between the two variables. Out of 11 sub-periods, seven witnessed movements in stock prices and oil prices in the same direction. Walayat’s explanation of positive correlation is that rising oil prices are associated with economic growth and improved corporate earnings, which should result in higher stock prices. On the other hand, some economists and observers contend that despite a Wall Street Journal headline like “Oil Spike Pummels Stock Market”, changes in oil prices do not affect stock prices in any predictable manner (Pescatori and Mowry, 2008). For example, sectoral factors play a role, as we should expect stock prices of oil companies to rise on higher oil prices, while the stock prices of financial institutions may rise when oil prices decline. In general the response of stock prices to changes in oil prices depends on whether the underlying firms belong to oil-related or oil-intensive sectors.

Intuitively, therefore, it is easier to explain the relation between stock prices and oil prices in oil exporting countries than in oil importing countries. Since there is as a good case for a positive as for a negative relation, this becomes an empirical issue. The objective of this paper is to produce further evidence by using the structural time series approach of Harvey (1989). This paper does make a contribution to the literature because Harvey’s technique makes it possible to account for variations in missing variables, which is a problem that has been largely ignored or overlooked.

2 Literature review

The evidence on the effect of oil prices on stock prices is mixed. Kling (1985), for example, concludes that rising oil prices are associated with stock market declines. Likewise, Papapetrou (2001) detects a negative oil prices effect on stock returns that extended over four months. Chen et al. (1986), in contrast, find that oil price changes have no effect on asset prices. Jones and Kaul (1996) report a stable negative relation between oil price changes and aggregate stock returns. Huang et al. (1996), however, fail to find a negative relation between stock returns and changes in the price of oil futures. Likewise, Wei (2003) concludes that the decline of US stock prices in 1974 cannot be explained by the 1973/1974 oil price increase. Evidence for a positive relation between the two variables has been produced by studies distinguishing between supply shocks and
demand shocks (for example, Killian and Park, 2009) and those based on observing the
two prices over time sub-periods (for example, Walayat, 2011). In general there is more
formal evidence for either a negative or no association than for a positive association.

Killian and Park (2009) argue that one limitation of these studies is that the price of
oil is considered to be exogenous with respect to the economy, which may not be a valid
assumption. For example, some economists have demonstrated that the price of oil
responds to the same economic forces that drive stock prices (for example, Barsky and
Killian, 2009). They identify as a second limitation of these studies the assumption that it
is possible to assess the impact of higher oil prices without knowing the underlying
causes of rising oil prices. It is plausible to suggest that different shocks in the oil market
have very different effects on the economy and on the price of oil, as has been
documented by Killian (2008, 2009), and that the relative importance of these shocks
evolves over time. Killian and Park contend that regressions relating stock returns to
innovations in the price of oil tend to be biased towards failure to find a significant
relation between stock prices and oil prices (see, for example, Sadorsky, 1999). They
show that the response of aggregate US stock returns may differ significantly, depending
on whether the increase in the price of oil is driven by demand or supply shocks. The
conventional wisdom that higher oil prices necessarily cause lower stock prices is shown
to apply only to oil market-specific demand shocks such as rising precautionary demand
for oil that reflect concerns about impending or anticipated shortfalls in future oil supply.
In contrast, positive shocks to the global demand for industrial commodities cause both
higher oil prices and stock prices.

Alsalman and Herrera (2013) bring attention to the distinction between symmetric
and asymmetric responses to positive and negative price innovations. They investigate
the effect of oil price innovations on the US stock market using a model that nests
symmetric and asymmetric responses to positive and negative oil price innovations. They
find no evidence of asymmetry for aggregate stock returns and only very limited
evidence for industry-level portfolios. Moreover, they point out that these asymmetries
do not match up well with the conventional views regarding energy-dependent sectors of
the economy. Instead, they suggest that asymmetries are more likely driven by the effect
of oil price innovations on expected and/or realised demand. They also find that positive
oil price innovations depress aggregate stock returns, as well as the returns of about 60%
of the industry-level stock returns.

Research in this area is sometimes classified into studies of developed countries and
those of developing and emerging countries. Studies of developed countries have been
conducted by Jones and Kaul (1996), Huang et al. (1996), Sadorsky (1999), Papapetrou
(2001), Ciner (2001), Yang and Bessler (2004), El-Sharif et al. (2005), Anoruo and
the other hand, Maghlyereh (2004), Onour (2007), Aliyu (2009), Nandha and
Hammoudeh (2007) and Narayan and Narayan (2010) explore the relation between oil
prices and stock prices in emerging and developing countries. Maghlyereh (2004)
examines the dynamic linkage between oil price and stock returns in 22 emerging
economies and reaches the conclusion that the oil market has little influence on stock
markets. Narayan and Narayan (2010) use the Johansen test and find evidence indicating
that oil prices, stock prices and exchange rates form a long-run relation (a cointegrating
vector) in Vietnam. They also show that both oil prices and exchange rates have a
positive and statistically significant effect on Vietnam’s stock prices in the long run, but
The effect of oil prices on stock prices

not in the short run. Given that the notorious Johansen test cannot produce robust results and that no one really knows what a cointegrating vector is, these findings should be handled with care (see, for example, Wickens, 1996). Actually, the reason why the results show a long-term effect is that the Johansen test is bound to produce at least one cointegrating vector, if not more, as the investigator wishes.

The results of studies dealing with the effect of oil prices on stock prices are not sensitive to the distinction between developed and emerging countries but rather between oil importing and oil exporting countries. Some economists argue that even if the negative relation between stock prices and oil prices holds in oil importing countries it does not hold for stock markets operating in oil exporting countries. Filis et al. (2011) investigate time-varying correlation between stock market prices and oil prices for oil importing and oil exporting countries. A DCC-GARCH-GJR approach is employed, using data from six countries: Canada, Mexico, Brazil, the USA, Germany and the Netherlands. The contemporaneous correlation results show that time-varying correlation does not differ for oil importing and oil exporting countries. They also find that correlation rises (falls) in response to aggregate demand-side oil price shocks, which are caused by macroeconomic fluctuations or world turmoil (for example, wars). On the other hand, they find that supply-side oil price shocks do not influence the relation between the two variables. The lagged correlation results show that oil prices exercise a negative effect on the stock market in all cases, regardless of the origin of the oil price shock. The only exception is what happened during the global financial crisis when lagged oil prices exhibited positive correlation with stock markets. Filis et al. conclude that in periods of significant economic turmoil the oil market is not a safe haven that offers protection against stock market losses.

3 Methodology

Related to the problems identified by Killian and Park (2009) is the problem of missing variables. We cannot plausibly suggest that stock prices or returns can be explained in terms of one variable only (oil prices), which is the implication of a two-variable regression. Since the stock price index is a macroeconomic variable, it has to be explained in terms of other macroeconomic variables, a list of which cannot be exhaustive. Hence structural time series modelling is used to account for missing variables without identifying them explicitly.

Another advantage of this methodology is that the model is estimated in a time-varying parametric framework, which is more appropriate. This is because there is no reason why we should assume that the relation between the two variables does not change over time. This point is actually raised by Pescatori and Mowry (2008) who suggest that ‘it is also possible that the relationship between oil prices and the stock market changes over time’.

A structural time series model relating stock prices to oil prices can be written as follows:

\[ s_t = \mu_t + \phi_t + \delta_t p_t + \epsilon_t \]  

where \( s \) is the logarithm of the stock price index and \( p \) is the logarithm of the price of oil. \( \mu_t, \phi_t \) and \( \epsilon_t \) are the time series components of \( s \): \( \mu_t \) is the trend, \( \phi_t \) is the cyclical
component and $\varepsilon_t$ is the random component. If oil price has any effect on stock prices, the coefficient $\delta$ must be statistically significant. The trend, which represents the long-term movement of the dependent variable, is represented by the general specification.

$$
\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{2}
$$

$$
\beta_t = \beta_{t-1} + \zeta_t \tag{3}
$$

where $\eta_t \sim NID(0,\sigma^2_\eta)$ and $\zeta_t \sim NID(0,\sigma^2_\zeta)$. The cyclical component is specified as:

$$
\phi_t = \rho(\phi_{t-1} \cos \theta + \phi^*_{t-1} \sin \theta) + \omega_t \tag{4}
$$

$$
\phi^*_t = \rho(-\phi_{t-1} \sin \theta + \phi^*_{t-1} \cos \theta) + \omega^*_t \tag{5}
$$

where $\phi^*_t$ appears by construction such that $\omega_t$ and $\omega^*_t$ are uncorrelated white noise disturbances with variances $\sigma^2_\omega$ and $\sigma^2_{\omega^*}$, respectively. The parameters $0 \leq \theta \leq \pi$ and $0 \leq \rho \leq 1$ are the frequency of the cycle and the damping factor on the amplitude respectively. The period of the cycle, which is the time taken by the cycle to go through its complete sequence of values, is $2\pi \theta$ (Harvey, 1989; Koopman et al., 2006).

The model is estimated in a TVP framework using maximum likelihood and the Kalman filter to update the state vector. If the coefficient on the explanatory variable turns out to be significant, while the trend and cycle are also significant, this means that the explanatory variable appearing explicitly on the right hand side of the equation is an important determinant of the dependent variable, but there are other determining variables whose effects are embodied in the trend and cycle. If these components are insignificant, then only the explanatory variable determines the dependent variable. And if the coefficient on the explanatory variable is insignificant while the trend and cycle (or one of them) is significant, then the dependent variable is determined by variables other than the one appearing explicitly in the equation.

The model is validated by examining its out-of-sample forecasting power using measures of forecasting accuracy that depend on the magnitude of the error and those that measure direction accuracy. These measures are calculated from the forecasting errors corresponding to a pair of actual and forecast values. The forecast value at time $t$ is

$$
\hat{S}_t = \exp(\tilde{S}_t) \tag{6}
$$

The mean absolute error and root mean square error are calculated (in percentage terms) as follows:

$$
MAE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{\hat{S}_t - S_t}{S_t} \right| \tag{7}
$$

$$
RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left( \frac{\hat{S}_t - S_t}{S_t} \right)^2} \tag{8}
$$
The effect of oil prices on stock prices

Theil’s inequality coefficient is calculated as the ratio of the RMSE of the model to that of the random walk. Hence

\[
U = \frac{1}{n-1} \sum_{t=1}^{n} \left( \frac{\hat{S}_{t+1} - S_{t+1}}{S_{t+1}} \right)^2
\]

Direction accuracy is calculated as:

\[
DA = \frac{1}{n-1} \sum_{t=1}^{n} a
\]

where

\[
a = \begin{cases} 
1 & \text{if } (\hat{S}_{t+1} - S_t)(S_{t+1} - S_t) > 0 \\
0 & \text{if } (\hat{S}_{t+1} - S_t)(S_{t+1} - S_t) < 0
\end{cases}
\]

DA, therefore, takes values between 0 and 1, such that a perfect model that predicts the direction of change correctly on each occasion has a DA of 1.

4 Data and results

The empirical results are based on a sample of monthly data covering the period January 2000–March 2014 (the USA, UK, Japan and Saudi Arabia) and the period January 2005–March 2014 (UAE, Qatar and Kuwait). Thus the sample covers three oil importers and four oil exporters. For the purpose of out-of-sample forecasting the forecasting period is April 2012–March 2014. The data were obtained from the International Financial Statistics of the International Monetary Fund.

The model estimation results are reported in Table 1, which displays the estimated components and coefficients of the final state vector as well as the coefficient of determination, \(R^2\), and the diagnostics for serial correlation (DW and Q), heteroscedasticity (H), normality (N) and predictive failure (PF). Q is the Ljung-Box statistic, which has a \(\chi^2\) distribution and H is a test statistic for heteroscedasticity with an \(F\) distribution. N is the test statistic for the normality of the residuals based on measures of skewness and kurtosis, while PF is the predictive failure test based on the same sample split as that used for the out-of-sample forecasting exercise.

The results show that the equations are well determined in terms of the goodness of fit measures and that they pass all of the diagnostic tests. The level of the trend is significant in all cases, but in no case is the cycle significant. Since the coefficient on the explanatory variable is significant in all cases, this means that oil prices explain the cyclical behaviour of stock prices rather well, but there are missing variables that affect the secular trend of stock prices. Since the coefficient on the explanatory variable is significantly positive in all cases, we infer that oil prices have a positive effect on stock prices in both oil exporting and oil importing countries.
Table 1  Model estimation results

<table>
<thead>
<tr>
<th>Location</th>
<th>USA</th>
<th>UK</th>
<th>Japan</th>
<th>UAE</th>
<th>Qatar</th>
<th>Saudi</th>
<th>Kuwait</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_t )</td>
<td>6.80</td>
<td>6.24</td>
<td>7.27</td>
<td>4.08</td>
<td>5.42</td>
<td>5.08</td>
<td>5.16</td>
</tr>
<tr>
<td>(41.47)</td>
<td>(36.36)</td>
<td>(42.14)</td>
<td>(7.90)</td>
<td>(15.56)</td>
<td>(21.56)</td>
<td>(17.26)</td>
<td></td>
</tr>
<tr>
<td>( \beta_t )</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>(1.02)</td>
<td>(-0.57)</td>
<td>(0.35)</td>
<td>(2.01)</td>
<td>(0.14)</td>
<td>(0.52)</td>
<td>(-0.59)</td>
<td></td>
</tr>
<tr>
<td>( \phi_t )</td>
<td>0.00</td>
<td>0.04</td>
<td>0.12</td>
<td>0.09</td>
<td>0.02</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td>(0.12)</td>
<td>(1.63)</td>
<td>(0.00)</td>
<td>(0.79)</td>
<td>(0.96)</td>
<td>(0.39)</td>
<td>(-0.51)</td>
<td></td>
</tr>
<tr>
<td>( \varphi_t )</td>
<td>0.00</td>
<td>-0.05</td>
<td>0.01</td>
<td>0.03</td>
<td>0.05</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>(0.00)</td>
<td>(-1.78)</td>
<td>(0.00)</td>
<td>(0.07)</td>
<td>(0.36)</td>
<td>(0.92)</td>
<td>(-0.12)</td>
<td></td>
</tr>
<tr>
<td>( \delta_t )</td>
<td>0.14</td>
<td>0.20</td>
<td>0.13</td>
<td>0.41</td>
<td>0.31</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td>(4.11)</td>
<td>(5.51)</td>
<td>(3.46)</td>
<td>(3.91)</td>
<td>(4.29)</td>
<td>(3.78)</td>
<td>(4.12)</td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.93</td>
<td>0.94</td>
<td>0.96</td>
<td>0.88</td>
<td>0.98</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The \( t \) statistics are placed in parentheses underneath the estimated coefficients. \( Q \sim \chi^2(8) \) for the USA, UK, Japan and Saudi Arabia and \( Q \sim \chi^2(5) \) for the others. \( H \sim F(56,56) \) for the USA, UK, Japan and Saudi Arabia and \( H \sim F(34,34) \) for the others. \( N \sim \chi^2(2) \) and \( PF \sim \chi^2(24) \).


The model does exceptionally well in out-of-sample forecasting. The results reported in Table 2 include the four measures of forecasting accuracy described earlier. The root mean square error is low and very close to (even numerically lower than) that of the random walk. In four out of the seven cases Theil’s inequality coefficient is numerically lower than one, which is very rare as it is rather difficult to outperform the random walk in out-of-sample forecasting (see, for example, Moosa, 2013). In terms of direction accuracy, the model predicts the direction of change in stock prices correctly on 78% of the occasions in one case (Saudi Arabia), which is a respectable level of forecasting accuracy. All in all it is a good model.

Table 2  Forecasting results

<table>
<thead>
<tr>
<th>Location</th>
<th>USA</th>
<th>UK</th>
<th>Japan</th>
<th>UAE</th>
<th>Qatar</th>
<th>Saudi</th>
<th>Kuwait</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>2.99</td>
<td>3.14</td>
<td>3.82</td>
<td>4.91</td>
<td>2.53</td>
<td>2.25</td>
<td>3.29</td>
</tr>
<tr>
<td>RMSE</td>
<td>3.55</td>
<td>4.24</td>
<td>4.80</td>
<td>6.36</td>
<td>3.27</td>
<td>2.85</td>
<td>4.08</td>
</tr>
<tr>
<td>U</td>
<td>0.88</td>
<td>0.86</td>
<td>1.02</td>
<td>0.99</td>
<td>1.11</td>
<td>0.95</td>
<td>1.17</td>
</tr>
<tr>
<td>DA</td>
<td>0.70</td>
<td>0.57</td>
<td>0.43</td>
<td>0.39</td>
<td>0.65</td>
<td>0.78</td>
<td>0.43</td>
</tr>
</tbody>
</table>
The effect of oil prices on stock prices

5 Concluding remarks

The finding that the relation between stock prices and oil prices is significantly positive for all seven countries, irrespective of whether they are oil exporters or oil importers, can be explained without difficulty. In the oil exporting countries examined in this paper, higher oil prices boost government revenue and hence expenditure, which is the main propelling force of the economy. The result is a positive effect on the stock market. In oil importing countries, rising oil prices are associated with strong economic activity and hence rising profitability and stock prices. Since both oil prices and stock prices are positively affected by economic activity, they appear to be positively correlated.

This finding should not be considered as an undisputable fact of life – after all this is economics, not physics. In economics, unlike physics, there are no laws, just empirical observations. For example, Pescatori and Mowry (2008) argue that the results may depend on the frequency of the data or the definition of the variables. There may also be some asymmetric effects in the sense that the results may depend on whether oil prices are rising or falling and whether they are high or low, which affects the elasticity of the demand for oil. Because of these caveats, the results of empirical studies should not be considered to represent the truth, the whole truth and nothing but the truth.

Acknowledgements

I am grateful to Kelly Burns for providing the data and excellent research assistance. I would like to thank two anonymous referees for useful comments.

References


**Notes**

1 It would perhaps be more appropriate to unify the sample size but this can be done only by reducing the sample for oil importers, which would represent some loss of information.

2 A question was raised by an anonymous referee about the possibility of multicollinearity arising from correlation between the trend and the explanatory variable. This is unlikely to be the case because the trend is not a variable but rather a time series component that is equivalent to the intercept term and a deterministic time trend in a conventional regression equation. In any case, when the correlation coefficient between the trend and the explanatory variable is calculated, it turns out to be statistically insignificant.

3 Earlier, a comment was made about the possibility of the results changing as a result of unifying the sample size for oil exporters and oil importers. To entertain this possibility, the model was re-estimated using a smaller sample for oil importers – the results turned out to be qualitatively similar to those reported in Table 1. The conclusion, therefore, stands.
Dynamic volatility spillover effect analysis between carbon market and crude oil market: a DCC-ICSS approach

Lean Yu, Jingjing Li and Ling Tang*

School of Economics and Management,
Beijing University of Chemical Technology,
Beijing 100029, China
Email: yulean@mail.buct.edu.cn
Email: tjijingjing@sina.com
Email: tangling@mail.buct.edu.cn
*Corresponding author

Abstract: This paper applies the Dynamic Conditional Correlation (DCC) model and Iterative Cumulative Sums of Squares (ICSS) model to investigate the volatility spillover effect between carbon emission market and crude oil market. In particular, an effective time-varying correlation analysis method, i.e. DCC, is first conducted to capture the dynamic linkage relationship between the two markets. Then, the ICSS method is used to explore the structural changes of such spillover effect and further identify the impacts of the related events on the linkage mechanism. Using the European Union Allowance (EUA) futures price and Brent crude oil futures price as study samples, some interesting findings can be obtained from the empirical study: (a) there exists an obvious positive relationship between the EUA and Brent markets; (b) such dynamic spillover effect varies with time and becomes somewhat smaller in Phase III than Phase II and (c) economic events (e.g., the financial crises) and political changes would structurally change the dynamic linkage mechanism.

Keywords: global energy issues; crude oil market; carbon emission market; volatility spillover; DCC; ICSS.


Biographical notes: Lean Yu is a Professor in School of Economics and Management at Beijing University of Chemical Technology (BUCT), Beijing, China. His research interests include artificial intelligence, computer simulation, decision support systems, knowledge management and financial forecasting. He has published 4 books and over 70 papers in journals including IEEE Transactions on Knowledge and Data Engineering, European Journal of Operational Research, International Journal of Intelligent Systems, and Computers & Operations Research. Dr. Yu is the Editor-in-Chief or a coeditor of several international journals.

Jingjing Li received the bachelor degree in applied physics from Tianjin University, Tianjin, China, in 2013. She is currently pursuing the master degree in management science and engineering in School of Economics and Management at Beijing University of Chemical Technology, Beijing, China. Her research interests include data mining and energy economy.
1 Introduction

To address the global climate change, Carbon Emissions Trading (CET) scheme has been widely considered as an effective carbon emissions mitigation tool. For example, the European Union Emissions Trading System (EU ETS) was launched in 2005 and has become the largest CET market all over the world (Liu and Chen, 2013), through three main phases: Phase I, lasting from the year 2005 to 2007, Phase II, lasting from 2008 to 2012 and Phase III, which tends to run from 2013 to 2020 (Reboredo, 2013). Similarly, China, as one of the largest developing countries, also raised seven CET pilots in 2011 and planned to extend it into a nationwide scale in 2015 (Cui et al., 2014). Like other commodities, the fluctuations of carbon price may be the main factor driving the carbon market and further determining the effect of CET in carbon emissions reduction. Therefore, investigating the fluctuation of carbon price has become an increasingly hot issue within the field of energy researches.

According to existing studies, carbon price has been proven to be influenced by various factors, e.g., institutional decisions, weather, events, macroeconomic and financial market and energy market (Alberola et al., 2008; Aatola et al., 2013). Amongst them, energy market, especially crude oil market, might be the most closely related to the fluctuations of carbon price (Alberola et al., 2008; Zhuang et al., 2014). For example, Peri and Baldi (2011) argued that there exist long-run nonlinear relationships between Brent and European Union Allowance (EUA) futures prices, by using the threshold co-integration method. Kanen (2006) pointed out that Brent prices are the main drivers of the natural gas prices, power prices and further carbon prices. Aatola et al. (2013) applied a set of metrological methods, e.g., Ordinary Least Squares (OLS) estimation, Vector Auto-Regression (VAR) and Granger causality test and discovered that the carbon and oil markets are positively correlated and the EUA price is a Granger cause of crude oil price. In contrast, Hammoudeh et al. (2014) adopted a quantile regression approach and found that an increase in crude oil price may generate a substantial drop in carbon prices when the latter is at a high-level. Zhuang et al. (2014) used the multifractal detrended cross-correlation analysis and proved that the return series of carbon and crude oil prices are significantly cross-correlated.

However, these above studies mainly focused on the general relationship during the whole sampling periods, failing to capture the time-varying linkage mechanism between the two markets. On the other hand, the dynamic volatility spillover effect is even more important to allow adjusting all related decisions and policies when the environment is changed with time (Sun et al., 2012). Under such a background, the Dynamic Conditional Correlation (DCC) model can be introduced as a quite promising model, which can effectively capture the time-varying dynamic relationships between different systems, while other analysis techniques currently used for the relationship between carbon and oil
markets can only provide the general results. In particular, the DCC model, developed by Engle (2002), can offer a robust analysis of time-varying linkages by allowing conditional asymmetries in both volatility series. Furthermore, it can effectively address the heteroscedasticity of the observed series data. Due to its unique merits, the DCC model has been widely applied in energy and financial area. For instance, Efimova and Serletis (2014) implemented the DCC model to study the dynamic correlations amongst energy commodity prices. Sadorsky (2012) employed DCC for the correlation analysis between the oil price and the stock prices of clean energy. Koch (2014) adopted DCC to explore the dynamic linkages amongst carbon, energy and financial markets. Dimitriou et al. (2013) applied the DCC model to analyse the correlation amongst five important emerging equity markets. Therefore, this study will utilise this powerful tool, the DCC algorithm, to analyse the dynamic linkage between carbon emission market and crude oil market.

In addition, for identifying the impacts of the related events on the linkage mechanism between carbon and crude oil markets, the structural breakpoint tests can be also employed as effective analysis tools for further understanding the dynamic volatility spillover effect. Actually, various structural breakpoint tests have been used as event study models for studying various markets, including energy markets. For example, Alberola et al. (2008) found that institutional decisions can fall into one notable factor affecting carbon prices during Phase I of the EU ETS, by applying a Chow breakpoint test method. Furthermore, by combining the structural change test with DCC, Lean and Teng (2013) used the Bai and Perron (2003) multiple structural breaks test to identify the breakpoints of dynamic correlation between different stock markets.

Generally speaking, this paper tries to investigate the dynamic correlation between the carbon and crude oil markets. In particular, the effective time-varying correlation analysis, i.e. DCC, is implemented to capture the dynamic relationship between the two markets. Furthermore, the structural changes of the spillover effect are investigated based on a competitive structure breakpoint test method, i.e. Iterative Cumulative Sums of Squares (ICSS) algorithm, for further identifying the impacts of the related events on the linkage mechanism. In empirical study, the dynamic correlations between the EUA and Brent crude oil markets are estimated to provide some helpful insights into CET design, especially for the developing countries.

The main aim of this paper is to investigate the dynamic volatility spillover effect between carbon and crude oil markets using the DCC-ICSS approach. The remaining part of this paper is organised as follows: Section 2 describes the DCC-ICSS methodology in detail. The empirical results for estimating the dynamic linkage mechanism between the EUA and Brent markets are reported in Section 3. Section 4 concludes the paper and outlines the future research direction.

2 Methodology

To investigate the dynamic linkage mechanism between the carbon and crude oil markets, an effective time-varying correlation analysis method, i.e. the DCC method is first implemented to capture the dynamic relationship between the two markets. And then a competitive structural breakpoint test method, i.e. the ICSS algorithm, is adopted to further identify the impacts of the related events on the linkage mechanism.
2.1 DCC

The DCC model developed by Engle (2002) is actually an extension of Conditional Correlation Coefficient (CCC) proposed by Bollerslev (1990). Compared to other estimation methods, DCC considers the influence of heteroscedasticity and introduces other explanatory variables into the mean equation and thus enhancing the rationality and generalisation of the model. Besides, DCC can effectively capture the dynamic correlation of multivariate data, with simple and direct estimation. Generally, there are two main steps in the DCC model, i.e. GARCH modelling and correlations estimation.

In the first step, a univariate GARCH model is built for each series data. The GARCH model contains a mean equation (i.e. AR(1)) and a variance equation (i.e. GARCH(1,1)). Given the \( i \)-th \( N \times 1 \) return vector, the mean equation and variance equation can be represented by:

\[
\begin{align*}
    r_{it} &= E(\tilde{r}_{i,t} | \Omega_{i,t-1}) + \epsilon_{it}, \quad \epsilon_{it} = \sqrt{h_{ii,t}} z_{it} \\
    h_{ii,t} &= \tau_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i h_{ii,t-1}
\end{align*}
\]

where \( r_{it} \) is the \( i \)-th observed return series in period \( t \), \( \Omega_{i,t} \) is the information set of the history data before the period \( t \) and \( \epsilon_{it} \) is the random error. The conditional mean in \( r_{it} \) is modelled in terms of AR(1). \( z_{it} \) is a \( N \times 1 \) random error vector that follows the normal distribution, i.e. \( E(z_{it}) = 0 \) and \( \text{Var}(z_{it}) = 1 \), \( h_{ii,t} \) is the time-varying conditional variance, \( \tau_i \) is a constant term, \( \alpha_i \) is the ARCH effect and \( \beta_i \) is the GARCH effect. Accordingly, a positive coefficient of \( \beta_i \) implies volatility clustering and persistency in the positive changes of a price return. The sum of \( \alpha_i \) and \( \beta_i \) indicates the persistency of the volatility shock.

In the second step, the dynamic correlation between the \( i \)-th and \( j \)-th return series is estimated. In this step, the covariance matrix \( H_t(N \times N) \) is first calculated by the following equation.

\[
H_t = D_t R_t D_t
\]

where \( D_t \) is a diagonal matrix with time-varying standard deviations on the diagonal in period \( t \) and \( R_t \) is the conditional correlation matrix.

\[
D_t = \text{diag}\left(\sqrt{h_{11,t}}, \ldots, \sqrt{h_{NN,t}}\right)
\]

\[
R_t = \text{diag}\left(q_{NN,-1/2}, \ldots, q_{NN,-1/2}\right)Q_t \text{diag}\left(q_{NN,-1/2}, \ldots, q_{NN,-1/2}\right)
\]

where \( Q_t(q_{ij}, t) \) is a \( N \times N \) symmetric positive definite matrix, representing the time-varying covariance matrix of residual.

\[
Q_t = (1 - \alpha - \beta)\tilde{Q} + \alpha \tilde{\xi}_{t-1} \tilde{\xi}_{t-1}^T + \beta Q_{t-1}
\]

where \( \tilde{Q} \) is the \( N \times N \) unconditional covariance matrix of the standardised residuals and \( \tilde{\xi}_t \) is the standardised residual terms. The coefficients \( \alpha \) and \( \beta \) are both nonnegative and satisfy \( \alpha + \beta < 1 \).
Based on the above two steps, the DCC between the \(i\)-th and \(j\)-th series at time \(t\) can be given by:

\[
\rho_{q_{ij}} = \frac{q_{ij}}{q_{ii}q_{jj}}
\]  

(7)

2.2 ICSS test

The ICSS test, proposed by Inclan and Tiao (1994), is a popular structure breakpoint test approach, based on cumulative sum (CUSUM) by Brown et al. (1975). The test aims to test whether and how any mutable break (or structural change) exists to structurally impact the observed series. In the ICSS algorithm, the target series is assumed to hold a stationary unconditional variance \(\sigma_1^2\) until a sudden change takes place at time point \(K_1\), the stationary unconditional variance then changes to another value \(\sigma_2^2\) until the next breakpoint \(K_2\) and such process repeats. To test a sudden break, a cumulative sum of square residuals is calculated by:

\[
C_k = \sum_{t=1}^{k} c_t^2; \quad k = (1,2,\ldots,N)
\]  

(8)

where \(c_t\) is uncorrelated random variables with zero mean and unconditional variance, which is generated by:

\[
c_t = \frac{y_t - \left(\sum_{i=0}^{t-1} y_i / t\right)}{\sqrt{\left(\frac{t+1}{t}\right)S_y^2}}; \quad t = (1,2,\ldots,N)
\]  

(9)

where \(S_y^2\) is the variance of total samples of time series data \(y_t\) \((t = 1,2,\ldots,N)\). The ICSS statistic can be defined as:

\[
D_k = \frac{C_k}{C_N} - \frac{k}{N}; \quad k = (1,2,\ldots,N); D_0 = D_N = 0
\]  

(10)

If there is no obvious breakpoint along the time series data, the statistic \(D_k\) always oscillates around zero; or otherwise, it significantly departs from zero.

3 Empirical study

The most important carbon market (i.e. the EUA market) and crude oil market (i.e. the Brent market) are focused in this study. Section 3.1 describes the sample data; Section 3.2 and Section 3.3, respectively, report the analysis results for the DCC estimation and ICSS test and Section 3.4 further discusses the results and presents some interesting findings.
3.1 Data

As for study samples, the largest and the most mature carbon market, i.e. the EU ETS and its nearest famous crude oil market, i.e. the Brent market, are especially focused in this study. The sample data are the daily EUA futures price and Brent futures price, which are originated from the Quandl website (http://www.quandl.com). The two series data cover the period from 8 April 2008 to 21 July 2014, with the sample size of 1583 observations. In particular, the data cover both Phase II and Phase III of the EU ETS. Besides, the large data scale can effectively avoid the inefficiency that the time-varying method, such as the DCC model, often suffer from in the case of small-scaled samples. In this paper, the futures prices of the carbon and crude oil are used due to both the better quality and high practicability for investors in setting risk management strategies (Reboredo, 2014). Figure 1 shows the original series of the EUA and Brent futures prices. From Figure 1, it can be easily found that the two series have some similar trends during a certain period of time, implying a potential linkage between them.

Figure 1  The original data of Brent and EUA futures prices from 8 April 2008 to 21 July 2014 (see online version for colours)

In practice, the return of each futures price is calculated for further analysis, i.e. \( r_t = \log(p_t) - \log(p_{t-1}) \), where \( p_t \) is the price at time \( t \). Figure 2 displays the return series of the two futures prices. For the return series of the Brent price, there is a period of centralised and drastic fluctuations, which lasted from 2008 to 2009. The hidden reason may lie in the famous Global Financial Crisis (GFC) which started from 2008. Meantime, the return series of EUA price fluctuated in a similar way, but with a much small influence amplitude. Moreover, there are two other special periods of centralised and drastic fluctuations for the EUA data. One is the period from June 2011 to June 2012 when the European Sovereign Debt Crisis (ESDC) was broken and the other is the period during the first half of 2013 when the EU ETS stepped into Phase III.

For clear explanation, Table 1 reports some basic descriptive statistics for the EUA and Brent futures price series and their return series. Compared to the price data, the statistics of the return series are quite small. For example, the mean of the return is close to zero. Compared to Brent returns, the negative value of the skewness statistic of the EUA returns suggests a greater probability of decreases in the EUA return. Both return series appear high values of the kurtosis statistic, which is consistent with the fat tails in
returns distributions. Meantime, the Jarque–Bera test strongly rejects the null hypothesis of normality of both return series. Moreover, the ADF test also strongly rejects the null hypothesis of existing an unit root, indicating the stationarity of the two return series. The serial correlation in the volatility of the returns series can be seen through the Ljung–Box statistic, which strongly rejects the null hypothesis. Finally, the Autoregressive Conditional Heteroskedasticity-Lagrange Multiplier (ARCH-LM) statistic indicates that ARCH effects are likely to be found in both the Brent and EUA returns series. Accordingly, all statistics of the return series support that the DCC model can be built for estimating their dynamic relationship. In addition, the positive values of linear and nonlinear correlation coefficients indicate a positive linkage between the two markets. However, the linkage seems not strong.

**Figure 2**  The return series of Brent and EUA futures prices from 9 April 2008 to 21 July 2014 (see online version for colours)

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Descriptive statistics of price series and return series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Brent</strong></td>
</tr>
<tr>
<td><strong>Panel A: statistics of price series</strong></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>89.96046</td>
</tr>
<tr>
<td>Median</td>
<td>86.83000</td>
</tr>
<tr>
<td>Max.</td>
<td>146.08000</td>
</tr>
<tr>
<td>Min.</td>
<td>36.61000</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>24.85771</td>
</tr>
</tbody>
</table>
### Dynamic volatility spillover effect analysis

**Table 1** Descriptive statistics of price series and return series (continued)

<table>
<thead>
<tr>
<th></th>
<th>Brent</th>
<th>EUA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel B: statistics of return series</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.00000</td>
<td>–0.00089</td>
</tr>
<tr>
<td>Median</td>
<td>0.00023</td>
<td>0.00000</td>
</tr>
<tr>
<td>Max.</td>
<td>0.12707</td>
<td>0.23923</td>
</tr>
<tr>
<td>Min.</td>
<td>–0.10946</td>
<td>–0.43474</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.02146</td>
<td>0.03458</td>
</tr>
<tr>
<td>Skewness</td>
<td>–0.08245</td>
<td>–1.03665</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>7.83137</td>
<td>22.74949</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>1540.43000***</td>
<td>25993.63000***</td>
</tr>
<tr>
<td>ADF</td>
<td>–42.20341***</td>
<td>–31.37129***</td>
</tr>
<tr>
<td>Q(20)</td>
<td>61.20610***</td>
<td>97.52950***</td>
</tr>
<tr>
<td>Q2(20)</td>
<td>2455.60000***</td>
<td>207.12000***</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>50.24632***</td>
<td>11.25234***</td>
</tr>
</tbody>
</table>

**Panel C: correlation of return series**

<table>
<thead>
<tr>
<th></th>
<th>Brent</th>
<th>EUA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson</td>
<td>0.18313***</td>
<td></td>
</tr>
<tr>
<td>Kendall</td>
<td>0.13578***</td>
<td></td>
</tr>
<tr>
<td>Spearman</td>
<td>0.20322***</td>
<td></td>
</tr>
</tbody>
</table>

Notes: ***, ** and * denote the result is statistically significant at 1%, 5% and 10% levels, respectively.

### 3.2 Results for DCC estimation

In order to explore the dynamic relationship between the two markets, the estimation of the bivariate AR(1)-GARCH(1)-DCC model between the Brent and EUA returns is conducted, as reported in Table 2. The parameters of ARCH and GARCH are estimated at 1% significant level and the values are non-negative, justifying the appropriateness of the GARCH specification. Furthermore, the freedom parameter \((df)\) is significant, confirming the rejection against Gaussian distribution. This finding also supports the results in Table 1. Similarly, the results also verify the choice of student distribution, i.e. \(t\)-distribution here, is an appropriate distribution. In addition, the average correlation between Brent and EUA is positively weak, even significant, in coincidence with the linear and non-linear correlation results, as listed in Table 1.

**Table 2** Estimates of AR(1)-GARCH(1,1)-DCC

<table>
<thead>
<tr>
<th></th>
<th>Brent</th>
<th>EUA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: estimates of AR(1)-GARCH(1,1)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)</td>
<td>–0.031609</td>
<td>0.025783</td>
</tr>
<tr>
<td>t-Stat.</td>
<td>–1.1590</td>
<td>0.8874</td>
</tr>
<tr>
<td>Const.(V)</td>
<td>0.013817</td>
<td>0.152241*</td>
</tr>
</tbody>
</table>
Table 2   Estimates of AR(1)-GARCH(1,1)-DCC (continued)

<table>
<thead>
<tr>
<th></th>
<th>Brent</th>
<th>EUA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: estimates of AR(1)-GARCH(1,1)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$-Stat.</td>
<td>0.1670</td>
<td>1.8820</td>
</tr>
<tr>
<td>ARCH</td>
<td>0.055482***</td>
<td>0.144697***</td>
</tr>
<tr>
<td>$t$-Stat.</td>
<td>0.0003</td>
<td>3.9460</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.940628***</td>
<td>0.854651***</td>
</tr>
<tr>
<td>$t$-Stat.</td>
<td>0.0000</td>
<td>25.5800</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel B: estimates of DCC model</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average CORR</td>
<td>0.191061***</td>
<td></td>
</tr>
<tr>
<td>$t$-Stat.</td>
<td>3.3710</td>
<td></td>
</tr>
<tr>
<td>Alpha</td>
<td>0.007336**</td>
<td></td>
</tr>
<tr>
<td>$t$-Stat.</td>
<td>2.2850</td>
<td></td>
</tr>
<tr>
<td>Beta</td>
<td>0.988293***</td>
<td></td>
</tr>
<tr>
<td>$t$-Stat.</td>
<td>176.4000</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>7.569171***</td>
<td></td>
</tr>
<tr>
<td>$t$-Stat.</td>
<td>9.2180</td>
<td></td>
</tr>
</tbody>
</table>

Notes: ***, ** and * denote the result is statistically significant at 1%, 5% and 10% levels, respectively.

For better understanding, the dynamic correlation coefficients between the Brent and EUA markets are shown in Figure 3. Generally, the linkage between the two markets appears positive and shows significant fluctuations, with the largest correlation coefficient of 0.38 on 24 November 2014 and the smallest of 0.05 on 20 December 2013. In order to identify the main change points of the dynamic correlation effect, ICSS will be used, as discussed in Section 3.3.

Figure 3   The dynamic correlation coefficients of Brent and EUA return series (see online version for colours)
3.3 Results for ICSS test

In this section, the ICSS algorithm is employed to discover the structural breakpoints in the dynamic correlation (i.e. the DCC series in Figure 3) of the Brent and EUA markets and the results are reported in Tables 3 and 4. All potential breakpoints tested at 5% significant level are listed in Table 3 and finally determined as the breakpoints at the significant level of 1% (as highlighted in Table 3 and listed in Table 4). In addition, the related events, including economic events and policy changing are found for analysis.

Table 3 The results of ICSS test

<table>
<thead>
<tr>
<th>Breakpoint</th>
<th>z-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1154</td>
<td>7.5980</td>
<td>0.0000</td>
</tr>
<tr>
<td>471</td>
<td>5.3168</td>
<td>0.0000</td>
</tr>
<tr>
<td>200</td>
<td>7.2295</td>
<td>0.0000</td>
</tr>
<tr>
<td>123</td>
<td>5.0405</td>
<td>0.0000</td>
</tr>
<tr>
<td>109</td>
<td>1.9998</td>
<td>0.0228</td>
</tr>
<tr>
<td>80</td>
<td>1.7610</td>
<td>0.0391</td>
</tr>
<tr>
<td>372</td>
<td>1.9315</td>
<td>0.0267</td>
</tr>
<tr>
<td>415</td>
<td>1.6697</td>
<td>0.0475</td>
</tr>
<tr>
<td>755</td>
<td>6.4015</td>
<td>0.0000</td>
</tr>
<tr>
<td>533</td>
<td>2.1565</td>
<td>0.0155</td>
</tr>
<tr>
<td>504</td>
<td>1.7142</td>
<td>0.0432</td>
</tr>
<tr>
<td>663</td>
<td>1.6799</td>
<td>0.0465</td>
</tr>
<tr>
<td>578</td>
<td>2.0388</td>
<td>0.0207</td>
</tr>
<tr>
<td>560</td>
<td>1.7074</td>
<td>0.0439</td>
</tr>
<tr>
<td>651</td>
<td>3.8709</td>
<td>0.0001</td>
</tr>
<tr>
<td>692</td>
<td>2.1851</td>
<td>0.0144</td>
</tr>
<tr>
<td>739</td>
<td>2.0148</td>
<td>0.0220</td>
</tr>
<tr>
<td>926</td>
<td>5.2826</td>
<td>0.0000</td>
</tr>
<tr>
<td>792</td>
<td>1.8675</td>
<td>0.0309</td>
</tr>
<tr>
<td>856</td>
<td>3.5511</td>
<td>0.0002</td>
</tr>
<tr>
<td>822</td>
<td>1.9764</td>
<td>0.0241</td>
</tr>
<tr>
<td>847</td>
<td>1.9104</td>
<td>0.0280</td>
</tr>
<tr>
<td>880</td>
<td>1.7209</td>
<td>0.0426</td>
</tr>
<tr>
<td>1067</td>
<td>4.7929</td>
<td>0.0000</td>
</tr>
<tr>
<td>1013</td>
<td>1.9222</td>
<td>0.0273</td>
</tr>
<tr>
<td>990</td>
<td>2.5224</td>
<td>0.0058</td>
</tr>
<tr>
<td>952</td>
<td>1.8164</td>
<td>0.0347</td>
</tr>
<tr>
<td>1111</td>
<td>1.8230</td>
<td>0.0341</td>
</tr>
<tr>
<td>1426</td>
<td>4.1006</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Table 3 lists the results of ICSS test and it can be easily seen that a total of 34 suspicious breakpoints can be discovered throughout the DCC series between the EUA market and Brent market, at 5% significant level. Furthermore, a total of 12 breakpoints are finally determined at 1% significant level, as highlighted in Table 3.

The breakpoints which are tested at 1% significant level are also listed in Table 4, together with their related impacting events. These breakpoints can be classified into seven types based on the main related events: the GFC, recovery from GFC, ESDC, delaying market revitalisation plan of carbon market, the monetary stimulus policies proposed by Japan and the EU and delaying carbon allowance releasing (advised and got through). In addition, the first three events can fall into financial crisis, which has already been proved to be the major factor structurally changing carbon price (Zhu et al., 2014), while the last four are policy changes in carbon markets.
3.4 Discussions

Combining the results of DCC estimation and ICSS test, some interesting findings can be obtained, as illustrated in Figure 4. It is noted that BP in Figure 4 is the abbreviation for the breakpoint.

**Figure 4** The integrated results of DCC estimation and ICSS test (see online version for colours)

Generally speaking, there exists an obvious positive relationship between the EUA and Brent markets, because all DCC values are larger than zero. However, such dynamic spillover effect varies with time and become somewhat smaller in Phase III than Phase II, as illustrated in Figure 4. One possible reason for this result may be that during Phase III, the carbon market has been undergoing a depression period with a quite low carbon price, failing in promoting the use of clean energy. However, the demand for crude oil holds an otherwise upward trend, stimulating the crude oil price. It is due to the different trends of the two markets, the dynamic linkage of EUA and Brent returns become increasingly weak, recently.

Moreover, the main trend of the dynamic spillover effect is largely impacted by some related important events. As for economic events, the results imply that during financial crises, the relationship effect between carbon and crude oil markets becomes increasingly larger, holding an upward trend. For example, during the GFC from 2007 to 2009, the DCC value between EUA and Brent returns significantly increased from 0.287 at BP 1 (i.e. 26 September 2008) to 0.328 at BP 2 (i.e. 16 January 2009). Similarly, during the toughest period of ESDC (from the year 2011 to 2012), the DCC value jumped from 0.146 at BP 5 (i.e. 7 March 2011) to 0.234 at BP 8 (i.e. 19 July 2012), by about 160%. The hidden reasons can be summarised as: the financial crises would have negative impact on the whole global economic system, significantly reducing the price of various commodities, including the two important energy commodities of carbon allowance and crude oil. For example, during GFC (from 2007 to 2009), the Brent futures price reduced from 103.54 to 46.57 dollars/barrel with a decrease rate of 55% and the EUA futures price similarly reduced from 26.17 to 12.98 eurocents/metric tonne with a decrease rate of 50%. Therefore, the same decreasing trends of carbon and crude oil prices lead to an increasingly strong relationship effect, during the financial crises.
As for policy changes, different policies may have different impacts on the spillover effect between carbon and crude oil markets from two aspects. On the one hand, the financial policies which would impact the whole global economic system (and thus the two target markets) may have a positive influence on the dynamic linkage between the two markets. For example, when the monetary stimulus policies was proposed by Japan and the EU in May 2013, the DCC value between the EUA and Brent returns increased from 0.101 at BP 10 (i.e. 1 May 2013) to 0.129 at BP 11 (i.e. 8 November 2013), with an increase rate of 122%.

On the other hand, the energy policies which only focus on one of the two markets may negatively impact the dynamic linkage between the two markets. For example, when delaying market revitalisation plan of carbon market was proposed in November 2012, the DCC value between the EUA and Brent returns declined from 0.167 at BP 9 (i.e. 19 November 2012) to 0.101 at BP 10 (i.e. 1 May 2013), with a decrease rate of 40%. The situation is the same when the plan of delaying carbon allowance releasing was advocated (BP 11, 8 November 2013) and got through (BP 12, 9 December 2013). The hidden reasons can be summarised into the following fact that since the policy only impacts one of the two markets, the price of the target market would be changed, whereas the other would remain the current situation. Therefore, such different response tends to significantly reduce the dynamic relationship effect between the two markets.

4 Conclusions

In this paper, we investigate the dynamic correlation between the EUA and Brent markets. In particular, an effective time-varying correlation analysis, i.e. DCC, is implemented to capture the dynamic relationship between the two markets, and then, a competitive structure breakpoint test, i.e. ICSS, is adopted to further identify the impacts of the related events on the linkage mechanism. By coupling the two effective methods, a novel DCC-ICSS approach is provided in this paper, which offers a new perspective for analysing the linkage between the carbon and crude oil markets.

In empirical study, the most important, largest and most mature carbon market (i.e. the EUA market) and its nearest crude oil market (i.e. the Brent market) are thoroughly investigated and some interesting findings can be obtained from the empirical results. Generally, there exists an obvious positive relationship between the EUA and Brent markets. Furthermore, such dynamic spillover effect varies with time and becomes somewhat smaller in Phase III than Phase II. In addition, economic events (e.g., the financial crises) and political changes would structurally change the linkage mechanism. The results can further offer helpful insights in making various decisions, e.g., investment, risk management and policy design. That is when studying carbon market or crude oil market, their dynamic relationship and the related impacting events of both economic events and political changes should be taken into account for decisions making and adjustment (Li et al., 2014).

4 Conclusions

In this paper, we investigate the dynamic correlation between the EUA and Brent markets. In particular, an effective time-varying correlation analysis, i.e. DCC, is implemented to capture the dynamic relationship between the two markets, and then, a competitive structure breakpoint test, i.e. ICSS, is adopted to further identify the impacts of the related events on the linkage mechanism. By coupling the two effective methods, a novel DCC-ICSS approach is provided in this paper, which offers a new perspective for analysing the linkage between the carbon and crude oil markets.

In empirical study, the most important, largest and most mature carbon market (i.e. the EUA market) and its nearest crude oil market (i.e. the Brent market) are thoroughly investigated and some interesting findings can be obtained from the empirical results. Generally, there exists an obvious positive relationship between the EUA and Brent markets. Furthermore, such dynamic spillover effect varies with time and becomes somewhat smaller in Phase III than Phase II. In addition, economic events (e.g., the financial crises) and political changes would structurally change the linkage mechanism. The results can further offer helpful insights in making various decisions, e.g., investment, risk management and policy design. That is when studying carbon market or crude oil market, their dynamic relationship and the related impacting events of both economic events and political changes should be taken into account for decisions making and adjustment (Li et al., 2014).

Besides crude oil market, the carbon market is also largely influenced by other various factors, and a much more comprehensive analysis can be conducted through thoroughly capturing the spillover effects amongst the carbon market and other different factors. Furthermore, besides dynamic relationship covering different time periods, the linkage mechanism can also be explored on different time scales under a multi-scale analysis approach (Tang et al., 2014a; Tang et al., 2014b; Yu et al., 2014). It is also
worth noticing that our empirical study only focuses the most mature carbon market of EUA and its nearest crude oil market of Brent. However, other carbon markets (e.g., China’s seven CET pilots) and crude oil markets (e.g., West Texas Intermediate (WTI)) should be also taken into account for identifying their similarities, differences and relationships. We will look into these issues in the near future.

Acknowledgements

This work is supported by grants from the National Science Fund for Distinguished Young Scholars (NSFC No. 71025005), the National Natural Science Foundation of China (NSFC Nos. 91224001, 71433001 and 71301006), National Program for Support of Top-Notch Young professionals and the Fundamental Research Funds for the Central Universities in BUCT.

References


Spillover effect of international crude oil market on tanker market

Yuying Yang and Chang Liu
Institute of Policy and Management,
Chinese Academy of Sciences,
Beijing 100190, China
and
University of Chinese Academy of Sciences,
Beijing 100049, China
Email: yyy876@126.com
Email: changl1002@163.com

Xiaolei Sun* and Jianping Li
Institute of Policy and Management,
Chinese Academy of Sciences,
Beijing 100190, China
Email: xlsun@casipm.ac.cn
Email: ljp@casipm.ac.cn
*Corresponding author

Abstract: The tanker shipping market has been regarded as a key extension of the global oil market and its uncertainty is related to the volatility of oil market. Therefore, for improving transportation risk management of oil shipping companies and oil imports countries, it is of significance to investigate the spillover effects across the oil market and the tanker market. Taking the volatility breaks into account, this paper establishes a VAR-BEKK-GARCH model for both the entire sample period from 1 June 2006 to 1 April 2014 and two sub-period samples. The empirical results provide evidence that the volatility of Brent market has more significant impacts on the tanker market than the WTI market in general. Moreover, the influence of oil markets on the tanker market during sub-sample 1 period (1 June 2006–23 April 2009) is stronger than that in sub-sample 2 (24 April 2009–1 April 2014).

Keywords: energy security; oil market; tanker shipping market; spillover effects; structural breaks; BEKK-GARCH.


Biographical notes: Yuying Yang received her BS in July 2010 from Zhengzhou University. She is now a PhD Candidate in the University of Chinese Academy of Sciences. Her major is Management Science and Engineering, and her interests focus on risk management and energy security.

Chang Liu is a Postgraduate Student in the University of Chinese Academy of Sciences. Her major is Management Science and Engineering, and her interests focus on risk management and energy economics.

Copyright © 2015 Inderscience Enterprises Ltd.
Xiaolei Sun is an Associate Professor in the Institute of Policy and Management, Chinese Academy of Sciences. She has a PhD from the University of Chinese Academy of Science. She has worked in two projects of the National Science Foundation of China. Her research interests focus on risk management, especially on credit risk.

Jianping Li is currently a Professor in Institute of Policy and Management, Chinese Academy of Sciences. He received the PhD degree in Management Sciences and Engineering from the University of Science and Technology of China, Hefei, in 2004. He is the Secretary and Executive Committee Member of International Academy of Information Technology and Quantitative Management (IAITDM). His research interests include credit risk assessment, operational risk modelling, country risk modelling and risk integration modelling.

1 Introduction

Oil is the paramount energy source and it is playing an increasingly dominant role in the global economy. According to BP’s Statistical Review of World Energy 2014, oil accounts for 32.9% of global energy consumption in 2013. Oil is also an important geopolitical resource and its distribution is extremely uneven. In particular, almost a half of the world’s proved reserves are concentrated in the Middle East, while the major oil-consuming countries are distributed in North America, Europe and Asia-Pacific, far away from the Middle East (BP, 2014). As a result, transportation systems are in great need to deliver oil from production locations to consumption locations. Internationally, there are three basic modes of transportation including pipeline, railway and marine transportation. In the light of the advantage in transportation cost, volumes, flexibility and so on, marine transportation is the major mode of oil transportation and 62% traded global oil is transported via sea.

As an important part of marine transportation and also a key extension of the international oil market, oil tanker market’s uncertainty is closely correlated to volatility of the oil market. It is generally acknowledged that international oil market affects the tanker market in two different channels and this is reflected in the links between oil price and tanker freight rates (Adland and Cullinane, 2006; Alizadeh and Nomikos, 2004; Glen and Martin, 2005; Poulakidas and Joutz, 2009). The first channel is that volatilities of the global supply in oil market have direct impacts on the demand for crude oil transport services. For instance, an unexpected increase in crude oil production causes an immediate increase in the demand for crude oil transport services. The second channel is that volatilities in crude oil prices have no direct impacts on the demand for crude oil transportation services but have direct impacts on shipping companies’ transportation costs. Bunker fuel is a product from crude oil and its price is related to crude oil price. Nottevoom and Vernimmen (2009) found that bunker prices constantly fluctuate due to market forces and crude oil price. As fuel cost is a considerable expense for a company in the tanker freight market, tanker operators would hope to have a higher freight rate to make up the raised transportation cost in facing a rising trend in crude oil price. As a result, the spillover effect of international oil markets on tanker market is also an important risk source that oil importing countries should take into consideration.
With rapid increase in oil imports in the past decades and its oil import dependence reaching 58% in 2013, China’s oil import risk has been a great concern (Yang et al., 2014; Li et al., 2014a; Li et al., 2014b). To improve oil import risk management, it is of significance to investigate the spillover effects of the international oil markets on tanker market. This paper contributes to the research on the relationship between oil market and oil transportation market from three aspects.

Firstly, this paper makes a detailed analysis on the volatility transmission mechanism between oil market and oil transportation market. At present, numerous studies have been dedicated to investigate the relationship between oil market and other economic variables (Cologni and Manera, 2008; Zhang et al., 2010; Wu et al., 2012; Guesmi and Fattoum, 2014; Efimova and Serletis, 2014). Few researches have dedicated to studying the relationship between oil market and tanker market and related researches just proved the existence of a relationship between freight rates and oil prices. For example, Alizadeh and Nomikos (2004) investigate the relationship between tanker freight rates, Brent and West Texas Intermediate (WTI) futures in respective trading routes and find evidence in favour of the existence of a long-run relationship between freight rates and oil prices. Glen and Martin (2005) suggest that forces of supply and demand can make the relationship between crude oil prices and spot tanker rates ambiguous. Poulakidas and Joutz (2009) analyse the impact that the spike in oil prices has on tanker rates using cointegration and Granger causality analysis. Sun et al. (2014) focuses on the multi-scale correlation between freight rates and oil prices and find that there are medium and long-term correlations between freight rates and oil prices.

Secondly, this paper makes further analysis on the spillover effect of oil markets on tanker market in different periods by combining with structural changes. It is well known that crude oil markets have often been subjected to infrequent structural changes due to economic, geopolitical extreme events, financial crisis, wars and natural disasters and so on. In the past decades, international crude oil prices have experienced great changes. In particular, during the period from 2006 to 2008, oil prices increased rapidly and peaked at a historic high level both in nominal and real terms at $144 per barrel (Brent) in July 2008. When global financial crisis broke out in 2008, crude oil price experienced a sharp slump in the period from August 2008 to January 2009; after the global financial crisis, crude oil price went back to increase slowly (BP, 2014). Volatilities of crude oil prices show different features in different periods with various driving forces. Consequently, neglecting structural breaks in the whole sample would induce biases in the spillover effect. In this paper, we allow for the possibility of structural breaks in modelling the conditional volatility.

Thirdly, by considering the difference between two largest international oil markets, both Brent and WTI oil market are selected to explore the spillover effects on tanker market. Brent and WTI are two major oil benchmarks in the world and the spread between Brent and WTI crude oil prices became larger since 2010, which shows the difference of two international oil markets in reflecting information changes. Spontaneously, it is exactly necessary to explore the different spillover effects of the two oil markets on tanker market in this paper.

Lots of methods for the volatility spillover effects have been proposed (Baele, 2005; Du et al., 2011) and multivariate Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models have been widely used to estimate the spillover effects in mean and/or volatility among different markets for their flexible features of characterising the variance–covariance of finance time series (Ewing et al., 2002; Li et al., 2009; Chang et al., 2011; Sun et al., 2012; Efimova and Serletis, 2014). Mensi
et al. (2014) point out that BEKK models capture the effects on the current conditional volatility of own innovations and lagged volatility as well as cross-markets shocks and the volatility transmission of other markets. More importantly, BEKK models provide further explanations of the origins, directions and transmission of intensity of the shocks in at least two markets. As a result, the BEKK-GARCH model is adopted in this paper to investigate the spillover effects of international oil markets on the tanker market.

The remainder of the paper is structured as follows. Section 2 introduces the methodology adopted in this paper. Section 3 describes the data and some preliminary analysis. Section 4 reports and discusses the empirical results. Finally, conclusions and directions for further research are given in Section 5.

2 Methodology

This section firstly introduces the framework of this study, and then the spillover models are presented, including mean spillover model (VAR) and volatility spillover model (BEKK-GARCH). Finally, the structural break detection method of the ICSS algorithm is presented.

2.1 The framework

In order to illustrate the main steps of our method clearly, the research framework is provided as follows.

To investigate the spillover effect between oil market and tanker market, the research is carried out on different samples, and the main steps are as follows.

1. The spillover effects are examined based on full sample data:
   - Step 1: VAR mean equation is adopted to investigate the mean spillover effect between oil markets and tanker market during the full sample period.
   - Step 2: Based on VAR models, a BEKK-GARCH model is established to compute covariance equations to investigate the volatility spillover effect between oil markets and tanker market during the full sample period.

2. The spillover effects are examined on subsample data. Accounting for breaks, we divide the full sample period into several sub-sample periods and examine the man–volatility spillover effect in every sub-sample period:
   - Step 1: ICSS algorithm is used to identify structural breaks, and then the full sample is divided into several sub-samples where structural breaks in volatility are founded.
   - Step 2: VAR mean equation is applied to explore the mean spillover effect between oil markets and tanker market in different sub-sample periods.
   - Step 3: Based on the VAR models, BEKK-GARCH models are established to compute covariance equations to investigate the volatility spillover effect between oil markets and tanker market in different sub-sample periods.
2.2 Spillover models

2.2.1 Mean spillover model

We assume that the conditional mean of returns on the oil markets and tanker market can be described by a Vector Autoregressive (VAR) model, proposed by Sims et al. (1990). VAR models have been used extensively in spillover effect studies. The VAR technique is very appropriate and popular because of its ability to characterise the dynamic structure of the model as well as its ability to avoid imposing excessive identifying restriction associated with different economic theories (Eltony and Al-Awadi, 2001; Cologni and Manera, 2008; Rahman and Serletis, 2011; Mensi et al., 2014). That is to say that VAR does not require any explicit economic theory to estimate the model. The use of VAR in macroeconomic variables can generate much empirical evidence. In a system with two variables, a VAR model can be set up as follows:

\[
\begin{align*}
    r_{1,t} &= \sum_{i=1}^{\rho} \theta_{2j} \cdot r_{2,t-i} + \sum_{j=1}^{\rho} \theta_{1j} \cdot r_{1,t-j} + \epsilon_{1,t} \\
    r_{2,t} &= \sum_{i=1}^{\rho} \theta_{1j} \cdot r_{1,t-i} + \sum_{j=1}^{\rho} \theta_{2j} \cdot r_{2,t-j} + \epsilon_{2,t}
\end{align*}
\]  

(1)
where \( r_{1,t} \) and \( r_{2,t} \) are the logarithmic returns of the oil prices and tanker freight rates, respectively. The parameter \( p \) is the lag-order. The residuals, \( \epsilon_{1,t} \) and \( \epsilon_{2,t} \), are assumed to be serially uncorrelated. The coefficients \( \theta_{1j} \) and \( \theta_{2j} \) provide the measures of own-mean spillover, whereas the coefficients \( \theta_{1i} \) and \( \theta_{2i} \) measure the cross-mean spillover between oil market and tanker market.

### 2.2.2 Volatility spillover model

Full BEKK-GARCH models, developed by Engle and Kroner (1995), are adopted to examine the volatility persistence of each market as well as cross-volatility spillover effects between markets. Generally, a full BEKK-GARCH model can be expressed as follows:

\[
\varepsilon_t | \psi_{t-1} \sim N(0, H_t) 
\]

\[
H_t = C C^* + \sum_{i=1}^{K} C C^* C C^* + \sum_{i=1}^{K} A A^* A A^* + \sum_{i=1}^{K} B B^* B B^* 
\]

where \( \psi_{t-1} \) is the information set containing all the information available up to time \((t-1)\), \( \varepsilon_t \) is the vector of residuals obtained from the VAR model (equation 1). \( C^* \), \( A^* \) and \( B^* \) are \( n \times n \) parameter matrices with \( C^* \) is triangular, \( A^* \) and \( B^* \) are \( J \times n \) parameter matrices, and the summation limit \( K \) determines the generality of the process. Especially, formula (3) will be positive define under very weak conditions. It will be shown to be a particularly convenient representation for estimation and for analysis of simultaneous equations systems.

According to Engle and Kroner (1995), there is an identification problem when \( K > 1 \) in the BEKK model. Therefore in this paper, we restrict that \( K = 1 \). So, the bivariate BEKK-GARCH model with \( K = 1 \) and without exogenous influences can be specified as follows:

\[
H_t = C C + A \varepsilon_{t-1} \varepsilon_{t-1}^* A + B H_{t-1} B 
\]

where \( C \) is a \( 2 \times 2 \) upper triangular matrix with three parameters and \( A \) is a \( 2 \times 2 \) square matrix of parameters and measures the extent to which conditional variances are correlated with past squared errors. \( B \) is also a \( 2 \times 2 \) square matrix of parameters, which depicts the extent to which current levels of conditional variances are related to past conditional variance. In our case, the total number of estimated parameters is 11. More specifically, the expression for the conditional variance for each equation can be expanded as:

\[
h_{11,t} = c_1^2 + a_1^2 \varepsilon_{1,t-1}^2 + 2a_1 a_2 \varepsilon_{1,t-1} \varepsilon_{2,t-1} + a_2^2 \varepsilon_{2,t-1}^2 + h_{11,t-1} + 2h_{12,t-1} h_{22,t-1} + h_{22,t-1} 
\]

\[
h_{22,t} = c_2^2 + a_2^2 \varepsilon_{2,t-1}^2 + 2a_2 a_3 \varepsilon_{2,t-1} \varepsilon_{3,t-1} + a_3^2 \varepsilon_{3,t-1}^2 + h_{22,t-1} + 2h_{21,t-1} h_{11,t-1} + h_{11,t-1} 
\]

In our case, \( h_{11,t} \) and \( h_{22,t} \) stand for the conditional variance of returns of the oil market and tanker market prices series, respectively. The coefficients \( a_{11} \) and \( a_{22} \) capture the ARCH effects, reflecting the impacts of own shocks or events of the oil market and
tanker market prices series, respectively, on volatility. The coefficients \( a_{12} \) and \( a_{21} \) capture the effects of cross-market shock interactions. Concretely speaking, \( a_{12} \) captures the effects of shocks or events of oil market on the volatility of the tanker market, and \( a_{21} \) is the exact opposite. The coefficients \( b_{11} \) and \( b_{22} \) capture the GARCH effects, reflecting the effects of own volatility persistence of the oil market and tanker market prices series, respectively. The coefficients \( b_{12} \) and \( b_{21} \) capture the volatility interactions between oil market and tanker market.

It is particularly noticeable that if \( a_{12} = a_{21} = b_{12} = b_{21} = 0 \), it can be concluded that conditional variance in oil returns or freight rates is only affected by its own shocks and volatility persistence and there is no spillover effects between two markets. Therefore, BEKK models provide further explanations of the origins, directions and transmission of the shocks in two markets.

### 2.3 ICSS algorithm

In order to investigate different spillover effects between oil market and tanker market during different periods, we take the structural change detections into consideration. The Iterated Cumulative Sums Of Squares (ICSS) algorithm developed by Inclan and Tiao (1994) is used to measure the number of (significant) sudden changes in variance in the time series, as well as to estimate the time point and magnitude of each detected sudden change in the variance. It is widely applicable in structural change detection for its simplicity and excellent performance. The algorithm assumes that the variance of a time series is stationary over an initial period of time, until a sudden change occurs because of a sequence of related events; the variance then reverts to stationarity until another market shock occurs. This process is repeated over time, generating a time series of observations with an unknown number of changes in the variance.

The ICSS algorithm is presented based on centred (and normalised) cumulative sum of squares, which is expressed as follows: given a series of independent observations from a normal distribution with zero mean and unconditional variance \( \sigma_t^2 \), \( \{r_t\} \), the cumulative sum of the squared observations from the start of the series to the \( k \)th point in time is expressed as \( C_k = \sum_{i=1}^{k} r_i^2 \), where \( k = 1, \ldots, T \); the centred cumulative sum of squares is expressed as \( D_k = \frac{C_k}{C_T} - \frac{k}{T} \), where \( C_T \) is the sum of the squared residuals from the whole sample period. The function \( D_k \) is used systematically to look for change points at different points of the series and more detail information can be found in the study of Inclan and Tiao (1994). Once the change points are identified using the ICSS algorithm, the periods of changes in volatility are analysed with potential factors.

### 3 Data

#### 3.1 Data description

Nowadays, two types of crude oil prices are widely used as the benchmark in oil pricing in the international oil market: Brent and WTI crude oil prices. Brent crude oil price has been used to assess up to 70% of the oil produced worldwide, especially in the European
and Asian regions. WTI is the recent main oil benchmark in the Americas. In this paper, daily spot prices of WTI and Brent crude oil prices are selected to reflect changes in different oil markets, which are sourced from the website of U.S. Energy Information Administration.

The tanker shipping market is described by two indices, the Baltic Dirty Tanker Index (BDTI) and the Baltic Clean Tanker Index (BCTI), published by the Baltic Exchange. The BDTI represents tanker routes of crude oil, while the BCTI covers tanker routes of oil derivatives (gasoline, benzene, etc.). The indices are defined as the sum of multiplications of the average rate for each route with the weighted factor of that particular route. Considering the focus on crude oil transportation, we select the BDTI as the benchmark of crude oil tanker freight rates.

The sample period of all above is from 1 June 2006 to 1 April 2014, and the number of usable observation is 1742. As shown in Figure 1, time series of the original Brent, WTI and BDTI exhibits a similar trend in some periods. For example, before 2010, the three series movements are in keeping with each other, and experience a peak point almost at the same moment in the mid-2008 and a lowest point in the end of 2009.

**Figure 2** Time series of the original Brent, WTI and BDTI (see online version for colours)

The continuous daily returns are calculated by taking the difference in the logarithms as follows:  

$$ r_{it} = 100 \times \frac{\ln p_{it} - \ln p_{i,t-1}}{\ln p_{i,t-1}} $$  

for each series. Figure 2 shows the dynamics of all return series. Upon close inspection, it appears that high volatilities are clustered in each series. Thus, it is possible that the series exhibits serial correlation of conditional variances. This observation is consistent with the statistical finding of excess kurtosis, suggesting that the variance of the series may be time varying. It can be found that returns of Brent and WTI exhibit a large fluctuation around 2009, indicating that compared with tanker market, oil market is more easily subject to the financial crisis.
Table 1 provides the descriptive statistics of the daily returns and the results of statistic tests. Table 1 shows that the skewness coefficients are positive for Brent and WTI and negative for BDTI. The kurtosis coefficients are above 3 for all the return series, indicating that the probability distributions of oil and tanker indices are skewness and leptokurtic. The normality is rejected according to the JB statistics for each series. The Q-statistics show that all the returns are serially correlated. The result of ADF unit root test shows that all the return series are stationary at the 1% level.

3.2 Structural breaks detection

The entire sample can be divided into several sub-sample periods based on the results of structural detection using the ICSS method. Furthermore, considering the significance of the oil market, we only detect the volatility breaks of Brent and WTI. Figure 3 shows the returns evidence of the Brent and WTI series with the points of structural change and ±3 standard deviations.

According to Figure 3, four breaks exist in both Brent and WTI oil returns and volatility breaks are found in both Brent and WTI oil returns in April 2009. So we select 23 April 2009 as a division point and two sub-samples periods are obtained. The first phase is from 1 June 2006 to 23 April 2009 and the second phase is from 24 April 2009 to 31 March 2014.
Figure 3  Returns series of Brent, WTI crude oil prices and BDTI index (see online version for colours)

Figure 4  Volatility breaks of return series of Brent and WTI crude oil prices (see online version for colours)
Table 2  Summary statistics of two sub-samples

<table>
<thead>
<tr>
<th></th>
<th>Sub-sample 1</th>
<th></th>
<th>Sub-sample 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Brent</td>
<td>WTI</td>
<td>BDTI</td>
<td>Brent</td>
</tr>
<tr>
<td>Max</td>
<td>22.03</td>
<td>33.76</td>
<td>12.38</td>
<td>7.22</td>
</tr>
<tr>
<td>Mean</td>
<td>–0.05</td>
<td>–0.05</td>
<td>–0.13</td>
<td>0.07</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>2.98</td>
<td>3.51</td>
<td>3.15</td>
<td>1.77</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.17</td>
<td>1.37</td>
<td>–2.58</td>
<td>–0.19</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>7.49</td>
<td>16.23</td>
<td>34.35</td>
<td>1.28</td>
</tr>
<tr>
<td>J-B</td>
<td>1528.2***</td>
<td>7354.90***</td>
<td>32779***</td>
<td>81.15***</td>
</tr>
<tr>
<td>(Q(10))</td>
<td>24.82</td>
<td>43.51***</td>
<td>308.32***</td>
<td>2.65</td>
</tr>
<tr>
<td>(Q(20))</td>
<td>43.49</td>
<td>69.01***</td>
<td>325.31***</td>
<td>12.76</td>
</tr>
<tr>
<td>(Q^2(10))</td>
<td>99.76***</td>
<td>148.84***</td>
<td>0.76</td>
<td>46.27***</td>
</tr>
<tr>
<td>(Q^2(20))</td>
<td>237.31***</td>
<td>189.70***</td>
<td>24.98</td>
<td>90.11***</td>
</tr>
<tr>
<td>obs.</td>
<td>652</td>
<td>652</td>
<td>652</td>
<td>1089</td>
</tr>
</tbody>
</table>

Notes: \( Q(n) \) and \( Q^2(n) \) refer to the statistics of the Ljung–Box test for autocorrelation of returns and the squared returns respectively. ** ***Significance at the 1% level.

Table 2 shows that the average of all returns are negative in sub-sample 1, while it is positive in sub-sample 2 as a result of recovering from the financial crisis. It also shows that all returns exhibit larger fluctuations in sub-sample 1, namely that the standard deviations of every return series are greater than that of sub-sample 2. The skewness coefficients are positive for oil returns and negative for BDTI in sub-sample 1, and all returns are negative in sub-sample 2. The kurtosis coefficients are above 3 for all the return series, except for the Brent and WTI returns in sub-sample 1, and all returns are negative in sub-sample 2. The \( Q \)-statistics show that all the returns or the squared returns are serially correlated. The result of ADF unit root test shows that all the return series are stationary at the 1% level.

4 Empirical results

This section firstly displays the estimates of the mean and volatility spillover effect of oil market on tanker market during the whole sample and makes a comparison between two oil markets. Then the results of different sub-sample periods are presented.

4.1 The whole sample analysis

The entire sample is firstly used to model the VAR mean spillover and then BEKK-GARCH volatility spillover model is established.
4.1.1 Mean spillover effect

Based on overall consideration of various criteria, such as Final Prediction Error (FPE), Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC) and Hannan–Quinn (HQ) information criterion, we choose 3 as the optimal lag order. Table 3 presents the estimates of VAR model for the entire sample.

<table>
<thead>
<tr>
<th></th>
<th>(r_{BDTI})</th>
<th>(r_{Brent})</th>
<th>(r_{BDTI})</th>
<th>(r_{WTI})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r_{BDTI}{1})</td>
<td>0.2583 (0.0240)**</td>
<td>-0.0315 (0.0222)</td>
<td>0.2586 (0.0241)**</td>
<td>-0.0158 (0.0254)</td>
</tr>
<tr>
<td>(r_{BDTI}{2})</td>
<td>0.1583 (0.0243)**</td>
<td>-0.0145 (0.0227)</td>
<td>0.1527 (0.0246)**</td>
<td>0.0199 (0.0259)</td>
</tr>
<tr>
<td>(r_{BDTI}{3})</td>
<td>0.0937 (0.0240)**</td>
<td>0.0417 (0.0222)*</td>
<td>0.0958 (0.0241)**</td>
<td>0.0488 (0.0254)*</td>
</tr>
<tr>
<td>(r_{Brent}{1})</td>
<td>0.0552 (0.0260)**</td>
<td>0.0055 (0.0241)</td>
<td>0.0011 (0.0230)</td>
<td>0.0079 (0.0242)</td>
</tr>
<tr>
<td>(r_{Brent}{2})</td>
<td>0.0965 (0.0261)**</td>
<td>-0.0233 (0.0241)</td>
<td>0.0277 (0.0229)</td>
<td>-0.0097 (0.0241)</td>
</tr>
<tr>
<td>(r_{Brent}{3})</td>
<td>0.0340 (0.0262)</td>
<td>-0.0228 (0.0242)</td>
<td>0.0619 (0.0229)**</td>
<td>-0.0481 (0.0241)**</td>
</tr>
<tr>
<td>(C)</td>
<td>-0.0197 (0.0596)</td>
<td>0.0257 (0.0552)</td>
<td>-0.0170 (0.0598)</td>
<td>0.0226 (0.0630)</td>
</tr>
</tbody>
</table>

WALD test

\(H_c: \theta_1 = \theta_2 = \theta_3 = 0\)  \(F = 6.6781***\)  \(H_c: \theta_1 = \theta_2 = 0\)  \(F = 2.9572**\)

\(H_c: \theta_2 = \theta_3 = 0\)  \(F = 1.6699\)  \(H_c: \theta_2 = \theta_3 = 0\)  \(F = 1.8392\)

Notes: The standard errors are given within parentheses.

***Significance at the 1% level.

**Significance at the 5% level.

*Significance at the 10% level.

As can be seen in Table 3 for the VAR, the BDTI returns are significantly influenced by previous BDTI returns, and the mean spillover is found from the returns of Brent and WTI crude oil price to oil tanker freight rates. Moreover, the impacts of Brent and WTI oil markets on oil tanker market are quite different. For the mean spillover of Brent oil market on oil tanker market, the coefficients of \(r_{BDTI}\{1\}\) and \(r_{BDTI}\{2\}\) are significant at the 5% and 1% levels, respectively. This indicates that the changes of Brent crude oil prices would impact the oil freight rates in two days, while for the WTI oil market, only the coefficients of \(r_{WTI}\{3\}\) is significant at the 1% level, indicating that the changes of WTI crude oil would impact the oil freight rates two days later.
Table 3 also provides the Wald test’s results for two different joint hypotheses regarding spillover effects. The first hypotheses is that the mean spillover effect of oil returns on BDTI returns does not exist, or the oil returns do not influence the BDTI returns, i.e. \( \theta_{11} = \theta_{12} = \theta_{13} = 0 \); the second hypotheses is that the mean spillover effect of BDTI returns on oil returns does not exist, i.e. \( \theta_{21} = \theta_{22} = \theta_{23} = 0 \). As the results of Wald tests shown in Table 3 indicate, there are mean spillover effects of both Brent and WTI oil returns on BDTI returns but no mean spillover effects from BDTI to oil markets and the spillover effects of Brent returns are stronger than that of WTI on BDTI returns.

### 4.1.2 Volatility spillover effect

The results of volatility spillover effects for the entire sample are presented in Table 4. According to equation (5), the major parameters including \( a_{11}, a_{22}, a_{12}, a_{21} \) and \( b_{11}, b_{22}, b_{12}, b_{21} \) influence the conditional variance at the square level and it is meaningless to discuss the positive and negative signs. So when analysing in the following parts, it refers to the absolute value.

**Table 4** Estimation results of the BEKK-GARCH model

<table>
<thead>
<tr>
<th></th>
<th>Brent-BDTI</th>
<th>WTI-BDTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_{11} )</td>
<td>0.1417 (0.0305)***</td>
<td>0.2664 (0.0452)***</td>
</tr>
<tr>
<td>( c_{21} )</td>
<td>-0.3877 (0.3387)</td>
<td>0.2509 (0.0491)***</td>
</tr>
<tr>
<td>( c_{22} )</td>
<td>0.5936 (0.2318)**</td>
<td>0.0000 (0.9999)</td>
</tr>
<tr>
<td>( a_{11} )</td>
<td>0.1699 (0.0150)***</td>
<td>0.2474 (0.0199)***</td>
</tr>
<tr>
<td>( a_{12} )</td>
<td>-0.3316 (0.0203)***</td>
<td>0.1125 (0.0183)***</td>
</tr>
<tr>
<td>( a_{21} )</td>
<td>-0.0186 (0.0230)</td>
<td>-0.0160 (0.0127)</td>
</tr>
<tr>
<td>( a_{22} )</td>
<td>0.9364 (0.0597)***</td>
<td>0.4086 (0.0314)***</td>
</tr>
<tr>
<td>( b_{11} )</td>
<td>0.9838 (0.0026)***</td>
<td>0.9643 (0.0063)***</td>
</tr>
<tr>
<td>( b_{12} )</td>
<td>0.0380 (0.0136)***</td>
<td>-0.0445 (0.0051)***</td>
</tr>
<tr>
<td>( b_{21} )</td>
<td>0.0062 (0.0087)</td>
<td>0.0047 (0.0033)</td>
</tr>
<tr>
<td>( b_{22} )</td>
<td>0.6723 (0.0234)***</td>
<td>0.9299 (0.0091)***</td>
</tr>
</tbody>
</table>

**WALD test:**

- \( H_0 : a_{12} = b_{12} = a_{25} = b_{25} = 0 \) (no spillover effect) \( W = 68.2839*** \) (no spillover effect) \( W = 83.9233 *** \)
- \( H_0 : a_{12} = b_{12} = 0 \) (no spillover effect of Brent on BDTI) \( W = 268.5566*** \) (no spillover effect of WTI on BDTI) \( W = 81.2011*** \)
- \( H_0 : a_{31} = b_{31} = 0 \) (no spillover effect of BDTI on Brent) \( W = 0.6929 \) (no spillover effect of BDTI on WTI) \( W = 2.0420 \)

**Notes:** Standard errors are given within parentheses.

***Significance at the 1% level.

**Significance at the 5% level.
According to Table 4, the parameters $a_{11}$ and $a_{22}$ are statistically significant, implying the presence of strong ARCH effects, which means that their own past shocks affect conditional variance in both oil market and tanker market. It is also found that the values of $a_{22}$ are greater than the values of $a_{11}$ throughout, indicating that the conditional variance of tanker market is more subject to its own past shocks than that of oil markets. The parameters $b_{11}$ and $b_{22}$ are also statistically significant, indicating that presence of strong GARCH effects, which means that their own past volatility affects significantly the conditional variance in both oil market and tanker market. Besides, the values of $b_{11}$ is close to one (0.9838 and 0.9656), indicating the evidence of strong persistence in oil markets’ conditional variance. Moreover, it is founded that the effects of GARCH are stronger than ARCH in both Brent oil market and WTI market, and the volatility of oil returns are more persistent than BDTI returns.

The parameters $a_{12}, b_{12}$ and $a_{21}, b_{21}$ capture the cross-volatility spillover between the two markets. The parameters $a_{12}, b_{12}$ are statistically significant, indicating that the conditional variance in BDTI returns is also affected by past shocks and volatility in oil markets. The parameter $a_{12}$ in the two models differs greatly, implying different oil markets have different spillover effects upon the tanker market. However, the parameters $a_{21}, b_{21}$ are not significant, indicating that no spillover effects exist from the tanker to the oil markets.

Table 4 also provides the Wald test results for three hypotheses regarding volatility spillover effects in every model. The results lead us to reject the first two null hypotheses at the 1% level and accept the last one, which again indicates the existence of spillover effects of oil markets on tanker market. However, there is no evidence of volatility spillover effects from tanker market to the oil market.

Additionally, it is found that volatility spillover from different oil markets into tanker market is quite different. In detail, the Brent oil market has a stronger influence upon the tanker market, which may be caused by multiple factors, especially the unique features of each oil market. The WTI oil price, quoted on New York Mercantile Exchange (NYME), mainly reflects the oil supply–demand condition in North America, whereas the Brent oil price, quoted on International Oil Exchange (ICE) in London, mainly reflects the supply–demand condition in European and Asian markets. Therefore, the different demand situations in oil market will cause changes in tanker market since the demand for crude oil transport services is derived from the imbalance between the supply and demand for crude oil. The rapid increase in oil demand in Asian emerging countries (such as China and India) will result in the increase of oil imports from the Middle-East and South Arica, especially the Brent oil imports, because Brent oil prices are usually taken as a benchmark in those countries’ traded oil. This may induce volatility spillover of the Brent oil market on oil tanker market. It is of significance for oil transportation companies and oil importers that the different market orientations between the two oil markets should be given sufficient consideration.

4.2 Sub-period analysis

In this part, the VAR mean spillover model and BEKK-GARCH volatility model are established for each sub-period.
4.2.1 Mean spillover effects

Based on the same criteria in Section 5.1.1, the optimal lag order chosen in two sub-period samples are 4 and 6, respectively. Tables 5 and 6 present the estimates of VAR model (equation 1). According to the parameters in Table 5, there is a bidirectional mean spillover between Brent and BDTI and the mean spillover effects from Brent oil returns into BDTI returns are stronger than the effects from BDTI returns on Brent oil returns in sub-sample 1, whereas there is only unidirectional mean spillover from Brent oil returns into BDTI returns in sub-sample 2.

Table 5 VAR model estimates (BDTI-Brent)

<table>
<thead>
<tr>
<th>Sub-sample 1</th>
<th>Sub-sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{BDTI} )</td>
<td>( r_{Brent} )</td>
</tr>
<tr>
<td>( r_{BDTI(1)} )</td>
<td>0.3535 ( (0.0399)^{**} )</td>
</tr>
<tr>
<td>( r_{BDTI(2)} )</td>
<td>0.1690 ( (0.0423)^{**} )</td>
</tr>
<tr>
<td>( r_{BDTI(3)} )</td>
<td>0.1070 ( (0.0429)^{*} )</td>
</tr>
<tr>
<td>( r_{BDTI(4)} )</td>
<td>0.0248 ( (0.0431) )</td>
</tr>
<tr>
<td>( r_{BDTI(5)} )</td>
<td>-0.1260 ( (0.0401)^{***} )</td>
</tr>
<tr>
<td>( r_{BDTI(6)} )</td>
<td>-0.0514 ( (0.0371)^{***} )</td>
</tr>
</tbody>
</table>

WALD test

<table>
<thead>
<tr>
<th>Sub-sample 1</th>
<th>Sub-sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0: \theta_{11} = \theta_{12} = \ldots = \theta_{15} = 0 ) (no mean spillover effect of Brent on BDTI)</td>
<td>( H_0: \theta_{11} = \theta_{12} = \ldots = \theta_{15} = 0 ) (no mean spillover effect of Brent on BDTI)</td>
</tr>
<tr>
<td>( W = 1.8941^{*} )</td>
<td>( W = 0.4813 )</td>
</tr>
</tbody>
</table>

Notes: Standard errors are reported in the parentheses.
***Significance at the 1% level.
**Significance at the 5% level.
*Significance at the 10% level.
The estimated results in Table 6 suggest that bidirectional mean spillover effects exist across WTI and BDTI in sub-sample 1 and the mean-spillover effects from WTI oil returns on BDTI returns are stronger than the effects from BDTI returns into WTI oil returns. This is consistent with the findings between Brent and BDTI returns. However, there are no spillover effects between WTI and BDTI in sub-sample 2; this is quite different from that of Brent oil market.

Table 6  VAR model estimates (BDTI-WTI)

<table>
<thead>
<tr>
<th></th>
<th>Sub–sample 1</th>
<th></th>
<th>Sub–sample 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sub–sample 1</td>
<td>Sub–sample 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sub–sample 1</td>
<td>Sub–sample 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{BDTI}^1$</td>
<td>0.3550</td>
<td>0.0048</td>
<td>$r_{BDTI}^1$</td>
<td>0.1515</td>
</tr>
<tr>
<td></td>
<td>(0.0402)**</td>
<td>(0.0508)</td>
<td></td>
<td>(0.0303)**</td>
</tr>
<tr>
<td>$r_{BDTI}^2$</td>
<td>0.1457</td>
<td>0.0046</td>
<td>$r_{BDTI}^2$</td>
<td>0.1172</td>
</tr>
<tr>
<td></td>
<td>(0.0427)**</td>
<td>(0.0540)</td>
<td></td>
<td>(0.0306)**</td>
</tr>
<tr>
<td>$r_{BDTI}^3$</td>
<td>0.1267</td>
<td>0.0751</td>
<td>$r_{BDTI}^3$</td>
<td>0.0774</td>
</tr>
<tr>
<td></td>
<td>(0.0432)**</td>
<td>(0.0546)</td>
<td></td>
<td>(0.0305)**</td>
</tr>
<tr>
<td>$r_{BDTI}^4$</td>
<td>0.0111</td>
<td>-0.0469</td>
<td>$r_{BDTI}^4$</td>
<td>0.1075</td>
</tr>
<tr>
<td></td>
<td>(0.0431)</td>
<td>(0.0545)</td>
<td></td>
<td>(0.0303)**</td>
</tr>
<tr>
<td>$r_{BDTI}^5$</td>
<td>0.0221</td>
<td>-0.0440</td>
<td>$r_{BDTI}^5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0427)</td>
<td>(0.0540)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{BDTI}^6$</td>
<td>-0.1307</td>
<td>0.1335</td>
<td>$r_{BDTI}^6$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0401)**</td>
<td>(0.0506)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{WTI}^1$</td>
<td>0.0222</td>
<td>-0.0148</td>
<td>$r_{WTI}^1$</td>
<td>0.0065</td>
</tr>
<tr>
<td></td>
<td>(0.0319)</td>
<td>(0.0403)</td>
<td></td>
<td>(0.0364)</td>
</tr>
<tr>
<td>$r_{WTI}^2$</td>
<td>0.0130</td>
<td>-0.0466</td>
<td>$r_{WTI}^2$</td>
<td>0.0705</td>
</tr>
<tr>
<td></td>
<td>(0.0315)</td>
<td>(0.0398)</td>
<td></td>
<td>(0.0363)*</td>
</tr>
<tr>
<td>$r_{WTI}^3$</td>
<td>0.1269</td>
<td>-0.0719</td>
<td>$r_{WTI}^3$</td>
<td>-0.0304</td>
</tr>
<tr>
<td></td>
<td>(0.0315)**</td>
<td>(0.0398)</td>
<td></td>
<td>(0.0364)</td>
</tr>
<tr>
<td>$r_{WTI}^4$</td>
<td>0.0562</td>
<td>-0.0812</td>
<td>$r_{WTI}^4$</td>
<td>0.0165</td>
</tr>
<tr>
<td></td>
<td>(0.0316)*</td>
<td>(0.0399)**</td>
<td></td>
<td>(0.0364)</td>
</tr>
<tr>
<td>$r_{WTI}^5$</td>
<td>0.0425</td>
<td>-0.1846</td>
<td>$r_{WTI}^5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0316)</td>
<td>(0.0400)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{WTI}^6$</td>
<td>-0.0145</td>
<td>-0.0113</td>
<td>$r_{WTI}^6$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0319)</td>
<td>(0.0404)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>-0.0527</td>
<td>-0.0551</td>
<td>C</td>
<td>0.0145</td>
</tr>
<tr>
<td></td>
<td>(0.1079)</td>
<td>(0.1363)</td>
<td></td>
<td>(0.0696)</td>
</tr>
<tr>
<td>WALD test</td>
<td>H$_0$: $\theta_1 = \theta_2 = \ldots = \theta_6 = 0$</td>
<td>$W = 3.6282^{***}$</td>
<td></td>
<td>H$_0$: $\theta_1 = \theta_2 = \ldots = \theta_4 = 0$</td>
</tr>
<tr>
<td></td>
<td>(no mean spillover effect of WTI on BDTI)</td>
<td></td>
<td></td>
<td>(no mean spillover effect of WTI on BDTI)</td>
</tr>
<tr>
<td></td>
<td>H$_0$: $\theta_1 = \theta_2 = \ldots = \theta_6 = 0$</td>
<td>$W = 1.7843^*$</td>
<td></td>
<td>H$_0$: $\theta_1 = \theta_2 = \ldots = \theta_4 = 0$</td>
</tr>
<tr>
<td></td>
<td>(no mean spillover effect of BDTI on WTI)</td>
<td></td>
<td></td>
<td>(no mean spillover effect of BDTI on WTI)</td>
</tr>
</tbody>
</table>

Notes:  Standard errors are reported in the parentheses.
***Significance at the 1% level.
**Significance at the 5% level.
*Significance at the 10% level.
4.2.2 Volatility spillover effects

Tables 7 and 8 present the estimates of the BEKK-GARCH model and Wald tests in two sub-samples. Table 7 shows that in sub-sample 1, the bidirectional volatility spillover effects exist across Brent and BDTI, while there is only a unidirectional volatility spillover from WTI to BDTI as parameter $a_{21}$ and $b_{21}$ are not statistically significant. Moreover, Wald tests exhibited in Table 8 show the identical results. In sub-sample 2, the volatility spillover effects between Brent returns and BDTI returns are the same as the spillover effects between the WTI and BDTI returns. That is, the conditional variances of both Brent and WTI markets are affected by their past own shocks and volatility, while the conditional variances of tanker market are affected only by own cross past shocks and volatility, but not affected by the past volatility original from oil markets. The results of Wald tests exhibited in Table 8 show that spillover effects exist from oil markets to tanker markets.

Table 7  
Estimation results of the BEKK-GARCH model

<table>
<thead>
<tr>
<th>Sub-sample 1</th>
<th>Sub-sample 2</th>
<th>Sub-sample 1</th>
<th>Sub-sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{11}$</td>
<td>0.1324 (0.0264)***</td>
<td>–0.1595 (0.0225)***</td>
<td>0.2410 (0.0307)***</td>
</tr>
<tr>
<td>$a_{12}$</td>
<td>–0.2302 (0.0243)***</td>
<td>–0.1778 (0.0308)***</td>
<td>–0.2272 (0.0291)***</td>
</tr>
<tr>
<td>$a_{21}$</td>
<td>0.0606 (0.0269)**</td>
<td>–0.0452 (0.0382)</td>
<td>0.0563 (0.0419)</td>
</tr>
<tr>
<td>$a_{22}$</td>
<td>0.3057 (0.0436)***</td>
<td>1.5401 (0.1042)***</td>
<td>–0.4311 (0.0623)***</td>
</tr>
<tr>
<td>$b_{11}$</td>
<td>0.9870 (0.0044)***</td>
<td>0.9812 (0.0054)***</td>
<td>0.9700 (0.0082)***</td>
</tr>
<tr>
<td>$b_{12}$</td>
<td>0.0230 (0.0084)***</td>
<td>–0.0085 (0.0400)</td>
<td>0.0312 (0.0125)**</td>
</tr>
<tr>
<td>$b_{21}$</td>
<td>–0.0199 (0.0086)**</td>
<td>0.0138 (0.0168)</td>
<td>0.0047 (0.0167)</td>
</tr>
<tr>
<td>$b_{22}$</td>
<td>0.9312 (0.0132)***</td>
<td>0.3560 (0.0797)***</td>
<td>0.8752 (0.0264)***</td>
</tr>
</tbody>
</table>

Notes: Standard errors are given within parentheses.
***Significance at the 1% level.
**Significance at the 5% level.

In Table 8, the value of coefficient $a_{12}$ in sub-sample 1 is greater than that in sample 2, indicating the tanker market is more influenced by the original past shocks from oil market in sub-sample 1 than that in sub-sample 2.

This may be attributed to the following causes. Firstly, in sub-sample 1, the volatilities in oil markets are mainly caused by unexpected demand of crude oil in emerging economies (Kilian, 2009; Kilian and Hicks, 2013), which are more likely to affect the demand of tanker transport services. Additionally, when a little supply capacity is available, the deteriorating relationship between supply and demand for crude oil transport services likely results in a higher freight rate since it needs time to invest in building new vessels or converting other vessels into crude oil tankers. On the other hand, due to the financial crisis, the economy has contracted from June 2008 to April 2009, which leads to a sharp drop in the demand for crude oil and tanker transport services. Therefore, the volatilities driven by the demand in crude oil markets have spillover effects on the tanker market in sub-sample 1.
### Table 8  Wald tests results

<table>
<thead>
<tr>
<th>Sub-sample 1</th>
<th>Sub-sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brent-BDTI</strong></td>
<td><strong>WTI-BDTI</strong></td>
</tr>
</tbody>
</table>
| $H_0 : a_{12} = a_{21} = b_{12} = b_{21} = 0$  
(no spillover effect) | $H_0 : a_{12} = a_{21} = b_{12} = b_{21} = 0$  
(no spillover effect) |
| $W = 3.7657^{***}$ | $W = 35.7244^{***}$ |
| $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of Brent on BDTI) | $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of Brent on BDTI) |
| $W = 3.4760^{***}$ | $W = 33.2955^{***}$ |
| $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of BDTI on Brent) | $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of BDTI on Brent) |
| $W = 5.7668^{*}$ | $W = 3.4161$ |
| $H_0 : a_{12} = b_{12} = a_{21} = b_{21} = 0$  
(no spillover effect) | $H_0 : a_{12} = b_{12} = a_{21} = b_{21} = 0$  
(no spillover effect) |
| $W = 70.8216^{***}$ | $W = 21.8465^{***}$ |
| $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of WTI on BDTI) | $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of WTI on BDTI) |
| $W = 67.0051^{***}$ | $W = 14.8024^{***}$ |
| $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of BDTI on WTI) | $H_0 : a_{12} = b_{12} = 0$  
(no spillover effect of BDTI on WTI) |
| $W = 3.5030$ | $W = 4.2258$ |

**Notes:**  
***Significance at the 1% level.  
*Significance at the 10% level.

Secondly, in sub-sample 2, the declining demand for oil in the economies such as Europe, America and Japan, combined the excess of capacity in tanker market, leads tanker freight rates to keep going down. The volatilities in crude oil prices caused by non-supply factors have no direct impacts on the demand for crude oil transportation services but have direct impacts on shipping companies’ transportation costs. Hence, the influence of volatilities in crude oil markets on tanker market is weaker in sub-sample 2. This finding is also supported by Shi et al. (2013), who suggest that crude oil non-supply shock does not significantly affect the crude tanker market. The character of dynamic changes of spillover effects between oil markets and tanker market is getting weaker, and in the near future, it is competition in tanker market that will bring about fluctuations in freight rates. These features should be highly valued by new shipbuilding market and the demolition market likewise.

## 5 Conclusions

Tanker shipping market has been regarded as a key extension of the global oil market and the volatility of oil freight rate poses a great challenge for oil shipping companies and oil importers to manage risks. This paper attempts to investigate the spillover effects across oil markets and tanker markets using a VAR-BEKK-GARCH approach. The main conclusions obtained from the empirical results are as follows.
The results of the VAR-BEKK-GARCH model provide evidence of spillover effects across crude oil markets and tanker shipping markets.

The unidirectional mean volatility spillover effects from oil market to tanker market are found in the whole sample period of 2006–2014. Surprisingly, the spillover effects from different oil markets into tanker market are quite different. The Brent oil market has stronger influence than WTI upon the tanker market, which can be attributed to their different roles in the international crude oil market.

Based on the change breaks, the entire sample is divided into two sub-samples. It is found that oil tanker market is more influenced by the oil markets in sub-sample 1 than that in sub-sample 2. The volatilities of oil markets have considerable influence on the tanker market in sub-sample 1, which is most probably due to unexpected demand for oil and the hardship of reallocating the vessels to reach a new equilibrium in the tanker market in a short time. On the other hand, the influence of crude oil market on tanker market is weaker in sub-sample 2. This is because the excess of capacity and the intense competition in tanker market during this period maybe the main driving factors of oil freight rates.

It is intriguing to extend this study in future research in at least two directions: (1) examining the asymmetric effects using the asymmetric GARCH model to investigate the spillover between crude oil markets and tanker market in face of positive and negative shocks; (2) considering all the volatility breaks and analysing spillover effects between oil markets and tanker market during every sub-sample.

Acknowledgements

The authors gratefully acknowledge the financial support from the National Natural Science Foundation of China (Nos. 71003091, 71373009 and 71133005), the programme of Youth Innovation Promotion Association supported by the Chinese Academy of Sciences (CAS) and the Key Research Programme of the Institute of Policy and Management, CAS. We wish to express our sincere gratitude to the anonymous referees for their constructive comments on an earlier draft of our paper, according to which we have improved the content.

References


Oil price and economic growth: an improved asymmetric co-integration approach

Jian Chai, Ying Yang* and Limin Xing

International Business School, Shaanxi Normal University, Shaanxi 710119, China
Email: chaijian0376@126.com
Email: yangyingwill@163.com
Email: xinglm0219@163.com
*Corresponding author

Abstract: Asymmetric co-integration is used in this paper to analyse the asymmetry relationship between oil price and economic growth. In contrast with widely used asymmetric co-integration method currently, three-way decomposition model is applied in this paper to explore more precise long-term asymmetric relationship between oil price and economy. Our empirical analysis not only concerns the China economy, but also includes the USA and Japan. The empirical results show that the asymmetric co-integration relationship between oil price and GDP cannot be proved in the three countries. However, according to the long-term relationship study, GDP is mostly effected by historically maximum oil price, especially in China and Japan.

Keywords: oil price; economic growth; asymmetric co-integration; three-way decomposition model; China; USA; Japan.


Biographical notes: Jian Chai is an Associate Professor at Shaanxi Normal University, and an assistant in Economic and Financial Department of National Center for Mathematics and Interdisciplinary Sciences. He is Chairman of 281st Young Scientists Forum of China Association for Science and Technology and International Conference on Business Intelligence and Financial Engineering (BIFE), Committee member of Youth Committee, Systems Engineering Society of China, Reviewer of the Energy policy (SSCI/SCI/EI), Sustainability Science (SSCI/SCI/EI), System Engineering–Theory & Practice, China Population, Resources and Environment etc. His main field is energy economy.

Ying Yang is a graduate student at International Business School of Shaanxi Normal University.

Limin Xing is a graduate student at International Business School of Shaanxi Normal University.
1 Introduction

Oil is known as the ‘blood of modern industrial’ for it is an important non-renewable energy. The ‘Blue Book of World Energy: Annual Development Report on World Energy (2013)’ estimated that China would become the largest oil consumer in the world. A lot of research shows that economic growth depends on oil consumption, which means oil price playing an important role on economic development. There are several reasons for us to analyse the effects of oil price change on economy in China, the USA and Japan. Firstly, USA, China and Japan are all economic power in the world and have a significant effect on other countries’ economy. Secondly, according to 2012 World Oil Consumption published by EIA, the oil consumption of USA was 18.5 millions of barrels per day, China was 9.9 and Japan was 4.7, which occupied top three of the largest oil consumption countries in the world. Finally, according to Top World Oil Net Importers in 2012, the USA imported 7381 thousand barrels per day; China imported 5502, while Japan imported 4591. So it is significance to understand the relationship between oil price and economy growth in these three countries.

Existing research on the relationship between oil price and economic growth does not reach an agreement. It is generally accepted that rising oil price will inhibit economic development (Hamilton, 1983; Kilian and Vigfusson, 2011; Narayan et al., 2014). While the relationship between oil price and economic growth has been proved non-linear and asymmetrical by some researchers, such as Mork (1989) pointed out that the influence of oil price rise is greater than the impact of oil price decrease on GDP. Hamilton (1996) defined the oil price rise as the Net Price of Oil Increment (NOPI), in order to distinguish the soaring oil price. Lardic and Mingnon (2006) applied the asymmetric co-integration method to explore the relationship between oil price and the economy. Zhang (2008) confirmed the existence of non-linearity between Oil shock and economic growth in Japan. Lardic and Mignon (2008) indicated that there is evidence for asymmetric co-integration between oil price and the USA economy, the G7, Europe and Euro area economies, based on asymmetric co-integration. However, according to Herrera et al. (2014), the hypothesis that the response of industrial production to oil price increases and decreases is asymmetric cannot be proved.

In contrast with previous literatures decomposed oil price and GDP series into rising and cutting two parts to discuss the asymmetric co-integration between oil price and economic growth, we decompose a series into three parts, cumulative historical maximum increases and cumulative sub-maximum increases and cumulative cuts, which help us to distinguish positive and negative increments of time series more specifically.

2 Asymmetric co-integration methodology

According to Gately (1992), time series can be decomposed into three parts: a cumulative maximum historical series, a cumulative recovery series and a cumulative cut series. We use the following three-way decomposition method for time series $X_t$:

$$ X_t = X_t^{++} + X_t^+ + X_t^- $$  \hspace{1cm} (1)
\[ X_{t}^{++} = \max\{X_{1}, X_{2}, \ldots, X_{t}\} \]
\[ X_{t}^{+} = \sum_{i=1}^{t} \max\{0, (X_{i}^{+} - X_{i-1}) - (X_{i}^{+} - X_{i})\} \]  
\[ X_{t}^{-} = \sum_{i=1}^{t} \min\{0, (X_{i}^{+} - X_{i-1}) - (X_{i}^{+} - X_{i})\} \]  

(2)

In this paper, the original time series \(X_t\) means oil price, \(X_{t}^{++}\) means the cumulative maximum historical price, \(X_{t}^{+}\) the cumulative recovery price, \(X_{t}^{-}\) the cumulative decreases of oil price and the GDP series can be decomposed as well.

Consider two integrated time series \(X_{1t}\) and \(X_{2t}\) and define \(X_{1t}^{+}\) and \(X_{2t}^{+}\) for \(j = 1, 2\), according to equations (1) and (2). Suppose that there exists a linear combination between \(X_{1t}^{++}\), \(X_{2t}^{++}\) and \(X_{2t}^{-}\) is as follows:

\[ Z_{t} = \beta_{0}X_{1t}^{+} + \beta_{1}X_{1t}^{-} + \beta_{2}X_{2t}^{+} + \beta_{3}X_{2t}^{-} + \beta_{4}X_{1t}^{-} + \beta_{5}X_{2t}^{-} \]  

(3)

Then, as stated by Schorderet (2004), \(X_{1t}\) and \(X_{2t}\) are asymmetrically or directionally co-integrated if there exists a vector \(\beta = (\beta_{0}, \beta_{1}, \beta_{2}, \beta_{3}, \beta_{4}, \beta_{5})\) with \(\beta_{0} + \beta_{1} \neq \beta_{2}\) or \(\beta_{3} + \beta_{4} \neq \beta_{5}\) (and \(\beta_{0} \neq 0\) or \(\beta_{2} \neq 0\) and \(\beta_{3} \neq 0\)) such that \(Z_{t}\) in equation (3) is a stationary process. The idea is that the relationship between variables might not be the same whenever they increase or decrease. To simplify the equation and without loss of generality, we suppose that only one component of each series appears in the co-integration relationship, as follows:

\[ Z_{1t} = X_{1t}^{++} - \beta^{++} X_{2t}^{-} \]
\[ \text{or} \quad Z_{2t} = X_{1t}^{+} - \beta^{+} X_{2t}^{-} \]
\[ \text{or} \quad Z_{3t} = X_{1t}^{-} + \beta^{-} X_{2t}^{-} \]

(4)

Because of the non-linear properties of \(Z_{jt}\), \(j = 1, 2\), estimation results of equation (4) are likely to be biased by ordinary least squares method with finite sample, afterwards, Schorderet (2004) suggests to estimate the auxiliary models by OLS:

\[ \varepsilon_{1t} = X_{1t}^{++} + \Delta X_{1t} - \beta^{++} X_{2t}^{-} \]
\[ \text{or} \quad \varepsilon_{2t} = X_{1t}^{+} + \Delta X_{1t} - \beta^{+} X_{2t}^{-} \]
\[ \text{or} \quad \varepsilon_{3t} = X_{1t}^{-} + \Delta X_{1t}^{+} + \Delta X_{2t}^{-} - \beta^{-} X_{2t}^{-} \]

(5)

As proved by West (1988), since the regressor has a linear time trend in mean, the OLS estimate of equation (5) is asymptotically normal and usual statistical inference can be done. In order to test the null hypothesis of no co-integration against the alternative of asymmetric co-integration, the traditional Engle and Granger procedure can be applied to equation (5).
3 Empirical analysis

3.1 Data

This article selected the quarterly data from 1992 to 2013 and Texas (WTI) crude oil spot price (POB) was chosen to represent the international crude oil price, real GDP was chosen to stand for economic growth. As GDP has an obvious seasonal trend, X12 method was used to adjust seasonally tend data. In order to minimise the possible heteroscedasticity and eliminate the multi-co-linearity, we take logarithm to deal with all the data.

Figure 1 shows the decomposition of WTI crude oil spot price. PMAX stands for cumulative increases of log of maximum historical oil price and PREC stands for cumulative recovery of log of oil price, PCUT stands for cumulative decreases of log of oil price. Firstly, through the PMAX curve, we can see that from the first quarter of 2004 to the third quarter of 2008, international oil price continuously rise and create record highs. By PREC curve, we can see that after reaching the highest historical price in the third quarter of 2008, oil price took on restorative growth and maintain small-scope fluctuation in high price from the third quarter of 2009 to the fourth quarter of 2013. And the breadth of V-shaped combined by PREC curve and PCUT curve, shows the activity level of oil price fluctuations.

Figure 1  Decomposition of the crude oil price (see online version for colours)

Notes: Data comes from the international monetary fund.

3.2 Unit root test

In case of spurious regression which caused by non-stationary, we apply Augmented Dickey–Fuller (ADF) test and Phillips–Perron (PP) test to verify the stationarity of time series at first. According to the results given in Table 1, all statistics cannot reject the null hypothesis in level series and all series are I (1) at the 1% significance level by PP test.
Results of the ADF test indicate that the level value of oil price and GDP of Japan reject the null hypothesis at 5% and 10% significance when model with trend and intercept. However, we can still seek common points while reserving difference from the ADF test and PP test, that all series are I (1).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Results of unit root test on individual series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level</td>
</tr>
<tr>
<td>P</td>
<td>−3.66(3)**</td>
</tr>
<tr>
<td>GDP(US)</td>
<td>−1.15(3)</td>
</tr>
<tr>
<td>GDP(CN)</td>
<td>−1.41(3)</td>
</tr>
<tr>
<td>GDP(GP)</td>
<td>−3.31(3)*</td>
</tr>
</tbody>
</table>

Notes: * (resp. **, *** ) means reject the null hypothesis at 10% (resp. 5%, 1%) significance level.
(1) Means model with none.
(2) Means model with intercept.
(3) Means model with trend and intercept.

Secondly, after GDP regression with oil price (include intercept), not only ADF and PP test, but also KPSS (Kwiatkowski-Philips-Schmidt-Shin) and Johansen co-integration test were applied to residual series. The results were shown in Table 2. According to ADF and PP tests, the residual series are stationary which means that oil price and GDP are co-integrated at 1% significance level in the three countries. However, the KPSS and Johansen test have different results. KPSS test indicate that oil price and GDP of the USA and China are co-integrated at 10% significance, while oil price and GDP are not co-integrated in Japan. The Johansen test shows that only GDP of China is co-integrated with oil price. It seems difficult for us to have a consensus opinion. In order to figure out the correct relationship between oil price and GDP in the USA, China and Japan, we need to go further to discuss if there exists an asymmetric relationship between these two series.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Results of unit root test on residual series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADF</td>
</tr>
<tr>
<td>GDP(US)</td>
<td>−2.80(1)***</td>
</tr>
<tr>
<td>GDP(CN)</td>
<td>−3.84(1)***</td>
</tr>
<tr>
<td>GDP(GP)</td>
<td>−2.86(2)***</td>
</tr>
</tbody>
</table>

Notes: * (resp. **, *** ) means reject the null hypothesis at the 10% (resp. 5%, 1%) significance level. The number of lags in the ADF regressions is given in parenthesis.

3.3 Asymmetric co-integration test

We apply three-way decomposition model to test the null hypothesis of co-integration, the following equation present the relationship between oil price and GDP.

$$LGDP^c_t + \Delta LGDP^r_t = \alpha^{++} + \beta^{++} LOIL^{++} + e_t$$ (6)
Oil price and economic growth

\[ LGDP_t^* + \Delta LGDP_t^* = \alpha^* + \beta^* \Delta OIL_t^* + \varepsilon_{1t} \]  

(7)

\[ LGDP_t^- + \Delta LGDP_t^+ + \Delta LGDP_t^{++} = \alpha^- + \beta^- \Delta OIL_t^- + \varepsilon_{3t} \]  

(8)

According to Schorderet (2004), we can apply the unit root test, which contains ADF, PP, KPSS and Johansen tests to residuals \( \varepsilon_{1t}, \varepsilon_{2t} \) and \( \varepsilon_{3t} \) to find out whether oil price and GDP are asymmetric co-integrated or not. We cannot depend on the normal critical values because of the estimated residuals are not original series. According to Lardic and Mignon (2008), we can use the critical values that given by Fuller (1976) for the ADF and PP test and for the KPSS and Johansen test, we use the critical values defined by Shin (1994) and Johansen and Juselius (1990).

Results are displayed in Table 3. According to all the tests, it appears that there is no evidence of co-integration between oil price and GDP in the USA, except for the KPSS test on \( \varepsilon_{1t} \). The null hypothesis of co-integration has been rejected by the PP test on \( \varepsilon_{2t} \) and \( \varepsilon_{3t} \), which prove the asymmetric co-integrated relationship between oil price and GDP in China, while the other tests did not. And oil price and GDP of Japan appear to be asymmetrically co-integrated according to the ADF and PP test on \( \varepsilon_{3t} \). Every unit root test methods have its own advantages and disadvantages. According to the unit root test on residuals, the evidences of asymmetric co-integration are not strong enough for us to conclude that the asymmetric co-integration relationship exists.

Table 3  Results of unit root test on residual series: tests for asymmetric co-integration

<table>
<thead>
<tr>
<th>Country</th>
<th>Residual</th>
<th>ADF</th>
<th>PP</th>
<th>KPSS</th>
<th>Johansen</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>( \varepsilon_{1t} )</td>
<td>-2.25(1)</td>
<td>-2.08</td>
<td>0.34*</td>
<td>14.74</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_{2t} )</td>
<td>-1.28(0)</td>
<td>-1.42</td>
<td>0.23</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_{3t} )</td>
<td>-2.04(1)</td>
<td>-2.59</td>
<td>0.29</td>
<td>6.2</td>
</tr>
<tr>
<td>CN</td>
<td>( \varepsilon_{1t} )</td>
<td>-2.66(2)</td>
<td>-2.24</td>
<td>0.29</td>
<td>7.66</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_{2t} )</td>
<td>-2.59(1)</td>
<td>-3.96**</td>
<td>0.25</td>
<td>5.23</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_{3t} )</td>
<td>-1.35(3)</td>
<td>-5.77***</td>
<td>0.28</td>
<td>7.66</td>
</tr>
<tr>
<td>JP</td>
<td>( \varepsilon_{1t} )</td>
<td>-2.57(0)</td>
<td>-2.66</td>
<td>0.28</td>
<td>13.05</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_{2t} )</td>
<td>-2.70(0)</td>
<td>-2.60</td>
<td>0.188</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon_{3t} )</td>
<td>-3.46*(0)</td>
<td>-3.46*</td>
<td>0.26</td>
<td>5.41</td>
</tr>
</tbody>
</table>

Notes: * (resp. **, *** ) means reject the null hypothesis at the 10% (resp. 5%, 1%) significance level. The number of lags in the ADF regressions is given into parenthesis.

Tables 4–6 display the regression results of equations (6)–(8) in sequence. The coefficients \( \beta^* \), \( \beta^- \) and \( \beta^- \) shows the long-term relationship between oil price and GDP in term of maximum historical, recoveries and cuts. In the USA, the effect of maximum historical oil price and GDP was the same as it shows in recoveries and cuts. However, as for China and Japan, the coefficient \( \beta^- \) is higher than \( \beta^- \) and \( \beta^- \), means that maximum historical oil price has larger impact on GDP than oil price recover and cut. It is noteworthy that \( \beta^- \) is similar to \( \beta^- \).
Table 4  Regression results of equation (6)

<table>
<thead>
<tr>
<th>Country</th>
<th>$\alpha^+$</th>
<th>$\beta^+$</th>
<th>Adj. $R^2$</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>0.17</td>
<td>0.20</td>
<td>0.78</td>
<td>0.08</td>
</tr>
<tr>
<td>CN</td>
<td>0.43</td>
<td>0.90</td>
<td>0.94</td>
<td>0.17</td>
</tr>
<tr>
<td>JP</td>
<td>0.02</td>
<td>0.08</td>
<td>0.90</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 5  Regression results of equation (7)

<table>
<thead>
<tr>
<th>Country</th>
<th>$\alpha^-$</th>
<th>$\beta^-$</th>
<th>Adj. $R^2$</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>–0.01</td>
<td>0.02</td>
<td>0.62</td>
<td>0.01</td>
</tr>
<tr>
<td>CN</td>
<td>–0.01</td>
<td>0.02</td>
<td>0.57</td>
<td>0.01</td>
</tr>
<tr>
<td>JP</td>
<td>–0.002</td>
<td>0.07</td>
<td>0.94</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 6  Regression results of equation (8)

<table>
<thead>
<tr>
<th>Country</th>
<th>$\alpha^-$</th>
<th>$\beta^-$</th>
<th>Adj. $R^2$</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>0.03</td>
<td>0.02</td>
<td>0.78</td>
<td>0.01</td>
</tr>
<tr>
<td>CN</td>
<td>0.04</td>
<td>0.02</td>
<td>0.48</td>
<td>0.02</td>
</tr>
<tr>
<td>JP</td>
<td>0.01</td>
<td>0.07</td>
<td>0.97</td>
<td>0.01</td>
</tr>
</tbody>
</table>

4 Conclusions

Based on the quarterly data from 1992 to 2013, this paper applies the three-way decomposition model to the asymmetric co-integration tests on oil price and GDP in China, the USA and Japan. The results of empirical analysis are as follows. According to the unit root test on the regression residuals, we have no enough evidence to reject the null hypothesis, which means the existence of asymmetric co-integration cannot be proved in the USA, China and Japan. However, the phenomenon of asymmetric relationship between oil price and GDP can be observed in their long-term relationship study, the historical maximum oil price has the largest impact on GDP, especially in China and Japan.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (NSFC) under Grant No.71103115.

References

Oil price and economic growth


Forecasting long-term and short-term crude oil price: a comparison of the predictive abilities of competing models

Zhongbao Zhou*, Ke Duan, Ling Lin and Qianying Jin

School of Business Administration, Hunan University, Changsha 410082, China
Email: ZhongbaoZhou@gmail.com
Email: Dkienne@hotmail.com
Email: LingLin2014@126.com
Email: Jqyhere@163.com
*Corresponding author

Abstract: In this paper, we apply several models to forecast the WTI monthly crude oil price from the long-term and short-term aspects. Then we use several diagnostic assessments to check the predictive abilities of the competing models. The results show that EGARCH model is more suitable for the short-term forecast, while the TARCH model is more appropriate for forecasting long-term oil price than other models.

Keywords: WTI crude oil price; price forecast; predictive ability; long-term forecast; short-term forecast.


Biographical notes: Zhongbao Zhou is a Professor in School of Business Administration, Hunan University, China. He received the Bachelor’s degree of applied mathematics, master’s degree of management science & engineering, PhD degree of control science & engineering from National University of Defense Technology, China.

Ke Duan is a Postgraduate Student in School of Business Administration, Hunan University, China.

Ling Lin is a PhD candidate in School of Business Administration, Hunan University, China.

Qianying Jin is a Postgraduate Student in School of Business Administration, Hunan University, China.
1 Introduction

Crude oil is of great significance to human beings. Due to its scarcity and non-renewable, it is reasonable that the crude oil price becomes a focus of the world. In strategic planning and project appraisals, the decisions made by producers and consumers can be totally different according to the crude oil price. For investors of crude oil markets, portfolio allocation and risk management also depend on the price. In addition, oil price plays an important role in all kinds of human activities. Therefore, crude oil price forecasting is of considerable interests for many researchers and government officers. Effective forecasting of crude oil price is a front-burner issue in energy economics.

Basically, existing approaches for forecasting crude oil price can be classified into two main categories: quantitative and qualitative approaches. For the former, econometric models are most frequently used in oil price forecasting. Pindyck (1999) pointed out the importance of unobservable state variables and successfully put them into the model and using a Kalman filter to estimate these variables. The results demonstrate that the model with a deterministic linear trend leads to more accurate forecasts. Radchenko (2005) extended the work of Pindyck (1999) by using an auto-regressive process in error terms instead of a white noise process. Wang et al. (2005) and Xie et al. (2006) forecasted the WTI crude oil prices by applying the ARIMA model. The results suggest that the ARIMA model performs fairly modest in forecasting. A considerable body of economic studies has been devoted to the GARCH class of models (e.g., ARCH, ARCH-M, GARCH, EGARCH, TGARCH, GARCH-M). For instance, Sadorsky (2006) adopted several statistical models such as TGARCH and GARCH to estimate daily crude oil price. Kang et al. (2009) compared the ability to predict volatility using various types of GARCH models. The results show that the fractionally integrated GARCH model is good for both Brent and Dubai. While the component GARCH model performs better for the WTI. Wei et al. (2010) extended the work of Kang et al. (2009) by utilising nine GARCH-type models. The out-of-sample forecasts indicate that it is brief of evidence that a single GARCH model outperforms the others. But it is clear that the linear GARCH-type models are better in short-run and the non-linear GARCH-type models perform better in long-run on predicting volatility. Burbidge and Harrison (1984) applied the VAR model into oil price prediction. The results show that there exists a relationship between oil price and numerous economic variables. Mirmirani and Li (2004) applied the VAR and ANN techniques to forecast the oil price movements. The analysis demonstrates that the BPN-GA model may outperform the VAR model.

Although there are numerous models for forecasting crude oil price, however, less attention is paid on the predictive ability of these models. Some researchers adopted some statistics to assess the predictive abilities of these competing models and find out which one is better at forecasting the crude oil price. Day and Lewis (1993) adopted several testing measures, such as Root Mean Squared Error (RMSE), Mean Error (ME), Mean Absolute Error (MAE) to evaluate the performance of predicting models. Sadorsky (2005) used some time series models to test the persistence on the WTI market, such as Simple Moving Average (SMA), Simple Exponential Smoothing (SES), Auto-Regressive (AR) and linear regression. Then the performance of these models is assessed by some index such as MSE, Mean Percentage Error (MPE), MAE, Mean Absolute Percentage Error (MAPE). Hansen (2005) proposed a comparable new method which is superior to
other techniques, such as Diebold and Mariano (DM) can prove to be more reliable in pairwise testing for two approaches. Patton (2011) utilised QLIKE and DM to rate the out of sample predictions and the measure was deemed to be robust.

This paper is to further explore the predictive abilities of a number of WTI crude oil prices’ forecasting models from the viewpoint of long-term and short-term. We adopt more performance criteria in order to compare these competing models more comprehensively. Some useful suggestions for choosing models are provided for long-term and short-term crude oil forecasting. The remainder of this paper is organised as follows. In Section 2, the raw data is processed to meet the requirements for the latter modelling. In Section 3, the models we adopted and performance measures to assess the predictive abilities of these competing models are introduced. The results are analysed in Section 4. Concluding remarks are summarised in Section 5.

2 Data

The price of crude oil has a positive impact on macro-economy, such as GDP, industry, agriculture and so on. Nowadays, there are three principal oil price markets all over the world, Brent (North Sea-European), Dubai (Persian Gulf) and West Texas Intermediate (WTI). The three markets are related to each other. Some researchers found that the crude oil price corresponds not only to the regional area but also global markets. In a sense, it would be essential for us to pay more attention to the oil price of volatility. In this paper, we adopt the monthly WTI crude oil price from January 2001 to July 2014 with 157 observations. Here we use monthly data instead of the daily or weekly data because of the consideration of seasonal factors. We plan to test the stability of the monthly data and try to catch the robust persistence and the long memory of competing models from the long-term and short-term aspects. The monthly data of WTI crude oil price is obtained from the Energy Information Administration (EIA). According to the previous literature, we define the real price of WTI crude oil as:

$$P_t = \frac{p_t}{ppi_t}$$

where $p_t$ and $ppi_t$ represent the WTI crude oil price and producer price index at time $t$, respectively.

As illustrated in Figure 1, the monthly crude oil price had experienced significant fluctuations during the fourteen years. With the unbalanced development of the global economy, regional attacks or terrorists shocked oil price worldwide, the price of crude oil dropped at the beginning of the year 2001 for the cause of 911 attacks. With the sophisticated demand-supply market, the unexpected oil demand surged the price in the following years of 2002–2006 since OPEC cutback its supply. Oil price fluctuation has closely related to the global economy which was reflected by the American credit crunch in the year of 2008, when it was the main extremely fluctuated in 14 years. Then it went down to the valley in the year 2009 for the OPEC cut the supply of the oil again. The rapid progress in emerging countries, global economic recovery and regional wars accelerated the demand of oil and the crude oil price increased in the year of 2010–2014.
Forecasting long-term and short-term crude oil price

Figure 1  Monthly WTI crude oil price and major events (June 2001–June 2014)

As shown in Figure 2, the sample of the mean and standard deviation of monthly WTI crude oil price are 0.398581 and 0.129701, respectively. A small skewness of −0.133087 implies that a larger probability of decreasing oil price may exist. A higher kurtosis of 2.398626 indicates that the price distribution has a fat-tail and intense peak with respect to the normal distribution. The Jarque-Bera of 2.829268 also dismisses the null hypothesis of normality.

Figure 2  The descriptive statistics of monthly WTI crude oil price (June 2001–June 2014)

Most of the time-series financial data reveals the non-stationary characteristics. Before decomposing the monthly data, we firstly use an ADF unit root test to test its seasonality. The results of ADF unit root test show that we should reject the null hypothesis under the significance levels of 1%, 5%, 10%. This indicates that the original series are non-stationary which contains the seasonality and trend.
Thus, we use X-12 seasonal adjustment approach to remove the seasonal factor at first. The results are shown in Figure 3. As revealed in Figure 3, most of oil price includes seasonal fluctuation. Influenced by seasonal factors, peak waves of oil price arrive in December every year. Seasonal fluctuations of oil price mainly present U-shape from 2002 to 2003 and 2004 to 2006. In 2007 and 2008, the fluctuations are relatively sharp because of financial crisis. After the year of 2009, the oil price has a stable increasing trend.

Figure 3 The trend of WTI crude oil price by the X-12 seasonal adjustment approach

Seasonal adjustment approach can eliminate seasonal factors from the time series, but the trend movements cannot be separated by the seasonal adjustment approach, thus we use HP filter approach to separate the long-term trend subsequently. The HP filter approach is described as follows.

Let \( \{ Y_t \} \) be trends and fluctuations in economic time series of factors, \( \{ Y_t^T \} \) be influence trend factor and \( \{ Y_t^C \} \) be fluctuation influence factor. Then we have:

\[
Y_t = Y_t^T + Y_t^C ; (t = 1, 2, ..., T)
\]

Using HP filtering operation can make the loss function minimised, that is

\[
\min \sum_{i=1}^{T} \left( Y_i - Y_i^T \right)^2 + \lambda \sum_{i=2}^{T} \left[ \left( Y_i^T - Y_{i-1}^T \right) - \left( Y_i^T - Y_{i-1}^T \right) \right]^2
\]

The first term is the deviation square and the second term is the trend component and two order difference square.

By using the above HP filtering approach, the trend of WTI crude oil price fluctuation is shown in Figure 4. After the two steps, we can see that the oil prices are not persistent and correlated with each other. Thus we should use some models to deal with the unstable oil price time series in order to catch the long persistence.
Forecasting long-term and short-term crude oil price

Figure 4 The trend of oil price fluctuation by the HP filtering approach and the cycle of WTI crude oil price by the X-12 seasonal adjustment approach

As illustrated in Figure 4, the left hand side indicates that the original WTI market crude oil price and the long-term trend of oil price. The right hand side of Figure 4 reveals the cycle and fluctuation after removing the seasonal factors. We can see that the oil price cycle is usually about one year. In this paper, we define the forecasting more than 12 periods as long-term forecast and the forecasting no more than 6 periods as short-term forecast. In order to compare the competing model from the viewpoints of long-term and short term WTI crude oil prices, we divide the samples into two parts, one for estimating and the other for test. The monthly data from July 2001 to December 2012 are used to estimate the parameters of forecasting models. Then the models are adopted to forecast the long-term crude oil price from January 2013 to July 2014 for testing their predictive abilities. Thus we use the monthly data from July 2001 to December 2013 to estimate the parameters of forecasting models. Then the models are used to forecast the short-term crude oil price from January 2014 to July 2014 for testing their predictive abilities.

3 Models and criteria

3.1 VAR model

A VAR model describes the evolution of a set of \( k \) variables over the same sample period \((t = 1, \ldots, T)\) as a linear function of only their past values. The variables are collected in a \( k \times 1 \) vector \( Y_t \).

Let \( Y_t = c + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \cdots + \Pi_p Y_{t-p} + \epsilon_t \), period \( t = 1, \ldots, T \), where \( \Pi_i \) is the \( n \times n \) coefficient matrix and \( \epsilon_t \) is an \( n \times 1 \) unobservable zero mean white noise vector process (serially uncorrelated or independent) with time invariant covariance matrix. The \( p \) order VAR is also called the VAR with \( p \) lags. The process of choosing the maximum lag \( p \) in the VAR model requires special attention because the inference is dependent on the correctness of the selection. In lag operator notation, the VAR \((p)\) is written as:

\[
\Pi(L)Y_t = c + \epsilon_t
\]

\[
\Pi(L) = I_n - \Pi_1 L - \cdots - \Pi_p L^p
\]
292  Z. Zhou et al.

\[ \det (I_s - \prod_{i=1}^{p} z \ldots - \prod_{i=1}^{p} z^r) = 0 \]  

(3)

\[ F = \begin{pmatrix} \prod_{i=1}^{p} & \prod_{i=2}^{p} & \ldots & \prod_{i=p}^{p} \\ I_s & 0 & \ldots & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & I_s & 0 \end{pmatrix} \]  

(4)

\[ \mu = (I_s - \prod_{i=1}^{p} z \ldots - \prod_{i=p}^{p})^{-1} c \]  

(5)

3.2 ARIMA model

Auto-Regressive Integrated Moving Average (ARIMA) model is a generalisation of an autoregressive Moving Average (ARMA) model. These models are mounted on time series data either to better understand the data or to predict future points in the series (forecasting). They are applied in some cases where data show evidence of non-stationarity, where a different initial step (corresponding to the ‘integrated’ part of the model) can be applied to remove the non-stationary. The model is generally referred to as an ARIMA\((p, d, q)\) model where parameters \(p\), \(d\) and \(q\) are non-negative integers that refer to the order of the autoregressive, integrated and moving average parts of the model, respectively. ARIMA models are an important component of the Box-Jenkins approach to time-series modelling. A ARMA\((p, q)\) model is given by:

\[ \left( 1 - \sum_{i=1}^{p} \phi_i L^i \right) X_t = \left( 1 + \sum_{i=1}^{q} \theta_i L^i \right) \xi_t \]  

(6)

where \(L\) is the lag operator, \(\phi_i\) are the parameters of the autoregressive part of the model, \(\theta_i\) are the parameters of the moving average part and the \(\xi_t\) are error terms. The error terms \(\xi_t\) are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean. Assume that the polynomial \(\left( 1 - \sum_{i=1}^{p} \phi_i L^i \right)\) has a unitary root of multiplicity \(d\). Then it can be rewritten as:

\[ \left( 1 - \sum_{i=1}^{p} \phi_i L^i \right) = \left( 1 - \sum_{i=1}^{p-d} \phi_i L^i (1-L)^d \right) \]  

(7)

An ARIMA \((p,d,q)\) process expresses this polynomial factorisation property with \(r=p-d\) and is given by:

\[ \left( 1 - \sum_{i=1}^{p} \phi_i L^i (1-L)^d \right) X_t = \left( 1 + \sum_{i=1}^{q} \theta_i L^i \right) \xi_t \]  

(8)

3.3 ARCH and TARCH models

Auto-Regressive Conditional Heteroskedasticity (ARCH) model modifies the assumption that the volatility of the time series is fixed. Let \(\epsilon_t\) denotes the return or residual return, it is assumed that \(\epsilon_t = \sigma_t z_t\), where \(z_t \sim iidN(0,1)\), thus the model can be written as:

\[ \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_p \epsilon_{t-p}^2 \]  

(9)
Threshold Auto-Regressive Conditional Heteroskedasticity (TARCH) model appears to be similar to the \textit{l-ARCH} model. TARCH\((m)\) model assumes the conditional variance was different from past observations, while the conditional variance was supposed to be evolved with previous residuals in TARCH model. TARCH\((m)\) model can be written as:

\[
X_t = \theta_1 X_{t-1}^+ + \theta_2 X_{t-1}^- + \cdots + \theta_m X_{t-1}^+ + \varepsilon_t
\]

\[
\varepsilon_t = \sigma_t \varepsilon_t
\]

\[
\sigma_t^2 = \alpha_0 + \alpha_1 \left( \varepsilon_{t-1}^+ \right)^2 + \alpha_2 \left( \varepsilon_{t-1}^- \right)^2 + \cdots + \alpha_m \left( \varepsilon_{t-m}^+ \right)^2 + \alpha_m \left( \varepsilon_{t-m}^- \right)^2
\]

where \(\varepsilon_t\) is an independent sequence with zero mean and unit variance, \(\varepsilon_t\) denotes the residue of price.

### 3.4 GARCH and EGARCH models

Generalised Auto-Regressive Conditional Heteroskedasticity (GARCH) model is the expansion of ARCH. The model is effective in excluding the excessive peak of the return on assets. Sadorsky (2006) showed that \textit{GARCH}(1,1) model performed well for describing the volatility of crude oil prices. The \textit{GARCH}(1,1) model can be written as:

\[
r_t = \mu + \varepsilon_t = \mu + \sigma_t z_t, \quad z_t \sim \text{NID}(0,1)
\]

\[
\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2
\]

where \(\sigma_t^2\) denotes the conditional variance.

Exponential Generalised Auto-Regressive Conditional Heteroskedasticity (EGARCH) model is similar to the GJR(1,1) model, that is \(\sigma_t^2 = \omega + [\alpha + \gamma I(\xi_{t-1} < 0)] \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2\), where \(I()\) is an indicator function and \(\gamma\) is the asymmetric leverage coefficient. EGARCH model may perform better than \textit{GARCH}(1,1) model, because the model allows the negative parameters \(\alpha\) and \(\beta\). The EGARCH model can be written as:

\[
\log(\sigma_t^2) = \omega + \alpha z_{t-1} + \gamma \left( |z_{t-1}| - E|z_{t-1}| \right) + \beta \log(\sigma_{t-1}^2)
\]

where \(\gamma\) is the asymmetric leverage coefficient used to describe the volatility leverage effect.

### 3.5 Forecasting criteria

In this paper, we adopt more criteria as follows in order to assess the predictive abilities of these above competing WTI crude oil price forecasting model comprehensively.

Suppose \(Y_t\) is the original series and \(\hat{Y}_t\) is the forecasted series, we define:

\[
e_t = Y_t - \hat{Y}_t
\]

Then the performance measures we adopted in this paper are defined as follows:

\[
ME = \frac{1}{T} \sum_{t=1}^{T} e_t
\]
\[ \text{MSE} = \frac{1}{T} \sum_{t=1}^{T} e_t^2 \]  
(14)

\[ \text{MAE} = \frac{1}{T} \sum_{t=1}^{T} |e_t| \]  
(15)

\[ \text{MMEU} = \frac{1}{T} \sum_{t=1}^{T} \begin{cases} |e_t| & \text{if } e_t < 0 \\ \sqrt{e_t^2} & \text{if } 0 \leq e_t \leq 1 \\ e_t^2 & \text{if } e_t > 1 \end{cases} \]  
(16)

\[ \text{MMEO} = \frac{1}{T} \sum_{t=1}^{T} \begin{cases} |e_t| & \text{if } e_t > 1 \\ \sqrt{e_t^2} & \text{if } -1 \leq e_t \leq 1 \\ e_t^2 & \text{if } e_t < -1 \end{cases} \]  
(17)

where \( T \) is the number of forecast preservation.

ME indicates the cost of errors is symmetrical. MSE penalises large errors more than small ones and the forecasters can assume whether small errors are preferable to a few large ones. In fact, MSE is a risk function and correspond to the expected value of squared error loss estimated by randomness for the forecaster who can’t possess large amount of information that can produce a more predictive estimate. MAE means that errors of the same magnitude are assigned the same weight regardless of their signs. MAE is regarded as the average over the real absolute value of the distinction between forecast and the corresponding observation. Besides, MAE is less sensitive to large errors than MSE. MMEU and MMEO are indicators of asymmetric measures that allow one to express his preferences by penalising under-predictions or over-predictions, especially large errors are penalised more heavily.

4 Empirical results

The long-term and short-term forecasting results of these models are shown in Figures 5 and 6, respectively. This paper assesses our oil returns volatility models based on various criteria. MSE, MAE, MMEU, MMEO have served as a measure to check the distant value between actual and forecast. ME is measures to verify whether the forecast model is underestimated or overestimated. The corresponding statistical measures related with these competing models are listed in Tables 1 and 2, respectively.

**Table 1** Performance measures of long-term forecasting of crude oil prices

<table>
<thead>
<tr>
<th>Measures and statistical tests</th>
<th>ARIMA</th>
<th>ARCH</th>
<th>TARCH</th>
<th>GARCH</th>
<th>EGARCH</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME</td>
<td>0.01652</td>
<td>-0.01873</td>
<td>0.00927</td>
<td>0.02701</td>
<td>0.01862</td>
<td>0.07018</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00068</td>
<td>0.00160</td>
<td>0.00064</td>
<td>0.00113</td>
<td>0.00079</td>
<td>0.00583</td>
</tr>
<tr>
<td>MAE</td>
<td>0.01861</td>
<td>0.03380</td>
<td>0.01844</td>
<td>0.02717</td>
<td>0.02061</td>
<td>0.07018</td>
</tr>
<tr>
<td>MMEU</td>
<td>0.11045</td>
<td>0.07124</td>
<td>0.08391</td>
<td>0.15142</td>
<td>0.12108</td>
<td>0.25769</td>
</tr>
<tr>
<td>MMEO</td>
<td>0.02978</td>
<td>0.13205</td>
<td>0.05334</td>
<td>0.02911</td>
<td>0.02973</td>
<td>0.07018</td>
</tr>
</tbody>
</table>
Forecasting long-term and short-term crude oil price

Table 2  Performance measures of short-term forecasting of crude oil prices

<table>
<thead>
<tr>
<th>Measures and statistical tests</th>
<th>ARIMA</th>
<th>ARCH</th>
<th>TARCH</th>
<th>GARCH</th>
<th>EGARCH</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME</td>
<td>0.00327</td>
<td>-0.00405</td>
<td>-0.00173</td>
<td>-0.00006</td>
<td>-0.00227</td>
<td>0.00389</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00060</td>
<td>0.00024</td>
<td>0.00023</td>
<td>0.00023</td>
<td>8.74E-05</td>
<td>0.00038</td>
</tr>
<tr>
<td>MAE</td>
<td>0.02063</td>
<td>0.01369</td>
<td>0.01266</td>
<td>0.01193</td>
<td>0.00684</td>
<td>0.01652</td>
</tr>
<tr>
<td>MMEU</td>
<td>0.07729</td>
<td>0.04541</td>
<td>0.04627</td>
<td>0.04679</td>
<td>0.02963</td>
<td>0.07711</td>
</tr>
<tr>
<td>MMEO</td>
<td>0.07969</td>
<td>0.08021</td>
<td>0.07094</td>
<td>0.06135</td>
<td>0.05215</td>
<td>0.06160</td>
</tr>
</tbody>
</table>

Figure 5  The long-term crude oil price forecasting with different models

Figure 6  The short-term crude oil price forecasting with different models
Table 1 and Figure 5 indicate that for the long-run, TARCH model fits best to support the forecast WTI crude oil price. It is clear that all the performance measures of the TARCH model are superior to others. EGARCH and ARIMA models present a stable and upward increasing predict direction for the cause of catching asymmetric leverage effect and decomposing seasonal element, respectively, while both of these models are less accurate to estimate the WTI crude oil price than the TARCH model for lack of consideration other disruptions. Table 1 shows that ME of the TARCH model is better than other models because its value is very close to zero, which means the cost of error tends to be symmetrical and the forecast error is very small. Although the ME of the ARIMA and EGARCH models are also very small, but they underestimate the WTI crude oil price and the results are less accurate with respect to the TARCH model. The MAE and MSE also show the similar results for the ARIMA and EGARCH models. MMEU and MMEO also indicate that the TARCH model outperform other models in forecasting long-term WTI crude oil price. The performance measures of the VAR and ARCH models show the deviations are larger than those of other models. Thus, it is not appropriate for the VAR and ARCH to forecast the long-term WTI crude oil price.

Table 2 and Figure 6 indicate that for the short-run, EGARCH model proves to be the best in predicting WTI crude oil price. Figure 6 reveals that the GARCH, TARCH, EGARCH, ARIMA, VAR almost have the same trend. Table 2 shows that the ME of the EGARCH model is close to zero, which means that the model is perfect in forecasting. ME of the EGARCH model is larger than that of the GARCH model because of the asymmetric effects, which means that the EGARCH model can allow the conditional variance to respond asymmetrically to negative shocks while GARCH model cannot persist this ability. ME, MAE and MSE of the EGARCH, GARCH, TARCH and VAR model are close to each other and show the similar trend in the short-term of crude oil price. But MMEU and MMEO indicate that the predictive ability of the EGARCH model is better than that of other models. It is evident that the performance criteria of the ARIMA model are inferior and the forecast is less accurate than other models, as shown in Table 2 and Figure 6. Thus, ARIMA is inappropriate to predict the short-term WTI crude oil price.

5 Conclusion

Forecasting crude oil price has a close relationship with our economy and other social activities. Our work focuses on the topic of assessing the predictive abilities of several competing forecasting models and seeking for the best technique which can apply for it. Previous studies show that there exist numerous available forecasting models. In this paper, we choose some linear and non-linear models to predict the WTI crude oil price from the long-term and short-term aspects. The empirical results show that the TARCH model is more suitable for long-term forecast, while the EGARCH model is more superior to the short-term forecast. The above results indicate that the WTI crude oil market may have more unknown factors in the long-run and short-run. Each model has its own characteristics and we should constantly revise and augment their variables to catch the rhythm of oil prices for better forecasting.
Acknowledgements

This research is supported by the National Natural Science Foundation of China (No. 71371067). We would like to thank Mr. Rencun Huang and Ms. Tiantian Ren for their useful suggestions.

References


The dynamics of crude oil price movements: from financialisation of crude oil prices and price expectation perspectives

Youngho Chang*, Shang Jyi Soon, Mirza Muhammad Hanif and Lucy Kusnadi

Nanyang Technological University, Singapore
Email: isyhchang@ntu.edu.sg
Email: sjsoon1@etu.edu.sg
Email: mirza1@e.ntu.edu.sg
Email: lkusnadi1@e.ntu.edu.sg
*Corresponding author

Abstract: This paper aims to examine the price pressure transmission effect from the crude oil futures market to the spot market through price expectation channel. It derives the relationship between the spot and the futures prices in the crude oil market and finds that the changes of the futures price and world output significantly predict the spot price movement. The impulse responses and the variance decomposition analyses demonstrate that the effect of the futures price on the spot price change is limited in a shorter period and diminishes as transmitting to the medium-run and the long-run while the effect of world output eventually increases and dominates in the long-run. The financialisation of crude oil prices indeed leads to a more drastic and frequent variation of price expectation among market participants that acts as a catalyst that destablises the crude oil prices.

Keywords: price expectation; VECM; spot price; futures price; financialisation.


Biographical notes: Youngho Chang is an Assistant Professor of Economics at the Division of Economics, Nanyang Technological University, Singapore. He is a member of Technical Committee for Clean Development Mechanism (CDM) Designated National Authority (DNA), National Environment Agency, Singapore. He specialises in the economics of climate change, energy and security, oil and macroeconomy, and the economics of electricity market deregulation. His current research interests are oil price fluctuations and macroeconomic performance, the economics of energy security, energy use and climate change, human capital and economic growth, and education for sustainability (EfS). He received his PhD in Economics from the University of Hawaii at Manoa, USA.

Shang Jyi Soon graduated from Nanyang Technological University with an honour bachelor degree in economics in 2014. During the course of study, he was qualified to participate in two campus-level undergraduate research
projects, such as URECA AY2012/2013 and FYP-URECA AY 2013/2014, under the supervision of Dr. Youngho Chang. The topic of former research project was ‘The Economics of Carbon Emissions Reductions’ while the latter’s was ‘The dynamics of crude oil price movements: from price expectation perspectives’. He earned the title of NTU Presidents’ Research Scholar as the result of these successful achievements.

Mirza Muhammad Hanif graduated from Nanyang Technological University with a double-major degree of bachelor with honours in economics and business in 2014. In his final year in university, he participated in the final year project under the supervision of Dr. Youngho Chang with the topic of ‘The dynamics of crude oil movements: from price expectation perspectives’.

Lucy Kusnadi was conferred the degree of bachelor with honours in Economics in 2014 associated with minor in Business and Entrepreneurship. While she was in university, she enrolled in the final year project with the topic ‘The dynamics of crude oil movements: from price expectation perspectives’ under the supervision of Dr. Youngho Chang.

1 Introduction

Noting the recent acceleration of crude oil financialisation associated with the drastic price movement, the study of crude oil price dynamics can no longer be limited within the scope of classical economic viewpoint whose model solely focuses on market fundamental factors. Before the financialisation of crude oil took place, any adverse fundamental shock, such as unusual shortages of supplies and extraordinary increased demands, would soon lead to the upward price movement. Otherwise, vice versa. Although the price response to any given market shocks still remains the same in terms of the direction in the 2000s where the trading of crude-oil-related financial derivatives are popular, it is apparent to observe that the magnitude and rate of price change has turned out to be significantly larger and different from that in the pre-2000 period (Carollo, 2012). A detailed illustration can refer to Figure 3. The phenomenon of such incredible price volatility has never been seen in the past despite the previous economic meltdowns and political shocks, such as: 1973 Oil Embargo, 1979 Iran Revolution, 1990 Gulf War, early 1980s recession and 1997–1998 Asian Financial Crisis, could have probably resulted in similar adverse impact to the crude oil market. The excess fluctuation of the crude oil price should be attributed to some ‘hidden’ emerging factors that have been overlooked and the classical supply and demand model seems to be too simple to predict the price change.

The study reviews the cross-market interaction between spot and futures prices based on the theoretical model of the price pressure transmission mechanism from a futures market to a spot market through a price expectation channel. The dynamic rational expectations of market participants that are subject to frequent changes of market conditions play a critical role in determining their trading behaviours and hence the market equilibrium. As such, the market price movement will never be steady but turbulent at all times, particularly in the context of a competitive crude oil market with high information accessibility and the low barrier of entry for market participants since
Therefore, it is reasonable to believe that market price equilibrium is highly affected by increasing dynamic price expectation which is determined by a wide range of complicating economic and financial behavioural factors.

The study constructs a theoretical model to explain the crude oil prices dynamics in both spot and futures markets. The model incorporates a variety of fundamental economic and financial factors to illustrate the frequent cross-market interactions based on the Rational Expectation Hypothesis. Beyond the foundation of related studies in the past, a theoretical mathematical function that represents price expectation is derived. With such, the study comes up with a hypothesis that the price expectation is very much dependent on the futures price as market participants tend to treat the futures price as a reliable price referencing indicator for investment and consumption decision making in practice.

There are a few notable contributions of the study to the literature. First, it integrates the equilibrium states of futures and spot markets to demonstrate the co-movements of futures and spot prices through a price expectation channel in the context of competitive global crude oil market. Second, the empirical analyses of this study provide a vivid picture about the interdependencies among the variables within a specific period and illustrate the responsiveness of each variable to another across the period.

This paper is organised as follows. Section 2 reviews empirical results and substantive findings of the related literature. Section 3 discusses an insight about the methodology, constructs a theoretical approach and provides the description of data employed for this study. Section 4 presents the results and interpretations of empirical analyses and highlight the significance of the results and the contribution of this study. Section 5 concludes this study.

2 The volatility and financialisation of crude oil prices

The financialisation of commodity prices to certain extent intensifies the price volatility. It has been shown that there is an increasingly positive correlation between the returns of non-oil commodities and the oil price in the 2000s. The effect of the price volatility from one commodity market on the price of its derivatives as a result of intense financial speculation could spill over to other commodity market. The speculative pressure in fact is contagious and may lead to herding behaviour among traders that pushes up the price higher than it would have been if it had been solely determined by fundamental factors (Tang and Xiong, 2012). Apart from the presence of price pressure transmission across markets of different types of commodities, Sanders and Irwin (2013) agree with ‘Masters Hypothesis’ that demonstrates that the price pressure is able to be transmitted from a single commodity futures price to its spot price through the arbitrage mechanism between these two markets. Any unprecedented buying pressure (or selling pressure) generated from index investment and commodities’ financial derivatives, say, futures, will sooner or later boost (or reduce) the futures price and hence spot price (Master and White, 2008).

The linkage between the futures and spot prices has been explained in a very quantitative way. Regarding the linkage, four propositions are derived from the study of a complex mathematical equation, incorporating most financial technical factors. First, investors’ expectations of economic activities in the future are as an important factor as
economic fundamentals in determining the spot price. Second, the changes of supply and demand in the futures and other types of financial derivatives markets of a commodity will exert pressure on its spot market and thus a spot price variation follows. Third, the variability of market risk premiums as a result of changes in investor’s willingness of bearing risk will eventually affect equilibrium in the spot and futures markets. Fourth, a convenience yield and a market price at the current period affect the equilibriums of the spot market, the futures market and other financial derivatives markets at the next period through varying precautionary demands (Singleton, 2011).

Apart from such a linkage through varying precautionary demands, a change in the futures price would directly lead to a co-movement of the spot price through the financial arbitrage channel. Lombardi and Robays (2011) construct a Structural Vector Autoregression (SVAR) model to evaluate the role of different kinds of shocks in determining the crude oil price. The study classifies market shocks into four types, namely a supply shock, an economic-activity-driven demand shock, an oil-specific demand shock and a destabilising financial shock. The forecast error variance decomposition analysis indicates that the destabilising financial shock affects the crude oil spot price movement significantly but just within a limited period. In the long-run, 90% of the spot price variation is explained by the change in market fundamentals while the remaining 10% is attributed to the price pressure transmission effect from the futures market.

Stoll and Whaley (2011) affirm that speculation activities in the commodity (including crude oil) futures market and a commodity index investment significantly contribute to the spot price dynamics of that commodity. The herding behaviour of longing (or shorting) the futures in the market will transmit upward (or downward) price pressure to the spot market of that commodity, either through a direct or an indirect channel. They also show the empirical evidence that there is a high correlation of returns between index commodity and non-commodity. The speculation event itself would indirectly exert its influence on commodities prices from one commodity futures market to another through the trading behaviour of uninformed and positive feedback traders.

There are some scepticism about the correlation between the spot price and the futures price. The correlation between the spot price and the futures price is insufficient to justify if the variation of the futures price predicts the spot price (Fattouh, 2012). A similar upward-trending price movement also appears in other types of commodities markets for which there are no futures exchanges and in which an index fund does not operate (Fattouh et al., 2012). It is plausible to generalise that the growing speculative activities in the futures market should be responsible for the upward surge of spot price of a commodity.

Fundamental factors remain intact in driving the spot price movement and the futures price volatility that depends on the trading hype in the financial market will inevitably influence the rapid fluctuation of the spot price through a price expectation channel. This study aims to justify such an argument in the following sections.

3 Model and data

This section is separated into two segments. First, a mathematical approach is adopted to derive the model equation of price expectation. Second, the empirical analysis that covers
the period in review from 1990 Q3 to 2013 Q3 is conducted to examine if the argument that price pressure transmission from the futures to the spot markets through a price expectation channel is valid.

There are several notable assumptions that are crucial to be recapitulated. First, market information is assumed to be scarce but valuable to investors. Second, the information accessibility is high to every market participant. Third, market participants are supposed to be rational enough to exploit useful information for investment decision making. Forth, all kinds of transactions in the markets are frictionless and costless. Finally, the market remains equilibrium such that the spot-futures spread is always zero so as to discard any arbitrage opportunity. Generally speaking, all these assumptions imply the presence of a constantly perfect and efficient market.

The equilibrium of a spot price is determined by the total quantity supplied and the total quantity demanded. The quantity supplied \( Q_{s,t} \) and the quantity demanded \( Q_{d,t} \), excluding inventory demand, are well represented by the following model equations, respectively.

\[
Q_{s,t} = \alpha_s + \beta_s P_t + \mu_s; \quad \text{where} \quad \mu_s \sim N(0, \sigma_s^2) \tag{1}
\]

\[
Q_{d,t} = \alpha_d + \beta_d P_t + \mu_d; \quad \text{where} \quad \mu_d \sim N(0, \sigma_d^2) \tag{2}
\]

The notation \( \alpha \) represents the intercept coefficient of both functions, \( \beta \) represents the slope coefficient which is also interpreted as the responsiveness of a quantity change to each additional unit increase of the price, \( \mu \) represents the residual term for each function and, theoretically speaking, is expected to be normally distributed so as to fulfill the requirement of best linear unbiased estimator (Muth, 1961; Gujarati and Porter, 2009). \( \mu_s \) incorporates other external factors, such as the number of new oil fields found, the entrance of new oil producers and the technological progress and efficiency improvement of crude oil exploration, drilling and production, which may cause a change of the quantity supplied of crude oil; meanwhile, \( \mu_d \) captures other external factors that affect the demand for crude oil; for example, they could be the demand for crude oil’s substitutes and complements, gross output and consumers’ preferences of using fuel products.

It is worth noting that the quantity supplied is a function of price expectation. This is in line with the industrial practice in reality that producers tend to make investment and production decisions based on their forecasts of price trends in advance before the market supply and demand clear at the equilibrium spot price. It is too perfect for the producers in such a competitive oil market to react instantaneously to any equilibrium price change since market clearing takes time to prevail. Producers always take an active and flexible role to review whatever market factors that are deemed influential to the spot price movement to adjust their quantities supplied. As such, according to Muth (1961), the quantity supplied at period \( t \) should be based on the producers’ price expectation that is determined by the lagged price at period \( t - 1 \) as follows:

\[
P_t^e = \theta_t P_{t-1} \quad \text{where} \quad \theta_t = f(\hat{\beta}_s, \hat{\beta}_d) \tag{3}
\]
The dynamics of crude oil price movements

The notation $\theta$, within a range from zero to one, is a time-varying parameter which measures the degree of sensitivity of price expectation at period $t$ to the lagged price at the preceding period $t - 1$. It is technically estimated by both demand elasticity and supply elasticity (Muth, 1961).

Inventory demand for crude oil or stockpile, $I_t$, constitutes a significant large part of the overall demand for crude oil. Therefore, the study defines the inventory demand separately that is represented by equation (4) rather than consolidates it with total market demand, $Q_{dt}$, as shown in equation (2). Commercial and speculative traders decide the capacity of crude oil inventory that they would like to hold based on their price expectations. Likewise, the price expectation is subject to a wide range of economics fundamental and financial factors. The inventory stockpile is basically determined by the difference between the expected price in next period $t + 1$ and the current spot price, say at period $t$.

$$I_t = \lambda \left( P_{t+1}^e - P_t \right) \quad (4)$$

$$I_t - I_{t-1} = \lambda \left[ \left( P_{t+1}^e - P_t^e \right) - \left( P_{t-1} - P_{t-1} \right) \right] \quad (5)$$

$$\Delta I_{t-1} = \lambda \left( \Delta P^e - \Delta P_{t-1} \right) \quad (6)$$

The notation $I_t$ represents the level of inventory stockpile at period $t$. $\lambda$ is a parameter, commonly positive, which measures the sensitivity of inventory demand to the expected price change. Equations (5) and (6) illustrate the change of levels of inventory stockpile, equivalently known as inventory demand, across time periods, a flow variable that is summed-up with the quantity demanded in equation (2) to form the total market quantity demanded.

In equilibrium state, the following condition is fulfilled, such that:

$$Q_{dt} = Q_{dj} + (I_t - I_{t-1}) \quad (7)$$

On the foundation of equation (7), a lengthy and tedious mathematical workout yields equation (8) that models the change of price expectation across time periods.

$$\Delta P_t^e = \frac{(\alpha_i - \alpha_d) + \beta_f P_t^e + \beta_s P_t + \lambda \left( P_t - P_{t-1} \right) + \left( \mu_{0,t} - \mu_{t, \Delta t} \right)}{\lambda} \quad (8)$$

The simplified version of equation (8) is as follows:

$$\Delta P_t^e = a + b_1 P_t^e + b_2 P_t + \Delta P_{t-1} \quad (9)$$

where $a = \frac{\alpha_i - \alpha_d}{\lambda}$, $b = \frac{\beta_f}{\lambda}$, $i \in \{s, d\}$, $u = \frac{\mu_{0,t} - \mu_{t, \Delta t}}{\lambda} \sim N(0, \sigma_u^2)$

In the futures market, the direction of futures price movement at each period depends on the total quantity of buying the futures contracts less the total quantity of selling them. It makes sense to summarise that the futures price at period $t$ is just a function of net trading positions of futures traders at that period.

$$F_t^Q = f \left[ \Delta \left( \text{Net Trading Position} \right) \right] \quad (10)$$
When the futures market is bullish, the more the long position exceeds the short position in terms of futures contract traded, the higher the futures price will grow. Otherwise, vice versa.

According to Lombardi and Robays (2011), futures price can be written in a simple equation form, such that:

$$F_{t+t}^Q = F_{t+t} + \delta_t$$  \hspace{1cm} (11)

In equation (11), $F_{t+t}^Q$ represents the quoted futures price with delivery date $\tau$ days from period $t$. $F_{t+t}$ represents the equilibrium futures price with delivery date $\tau$ days from period $t$ in the absence of arbitrage opportunity. It is also known as a fundamentally justified price level (Lombardi and Robays, 2011). $\delta_t$ represents the weakly stationary deviation of the quoted futures price from its equilibrium. In other words, it is also interpreted as the destabilising financial shock in the futures market. The shock can be attributed to a variety of unanticipated financial factors changes in the futures market. For instance, an unprecedented enormous increase of the trading volume (buying pressure) as a result of speculative hype in the futures market will instantaneously push the futures price up to a higher level that exceeds its fundamental fair value. However, such an increase of the futures price will not sustain as the arbitrage opportunity will eventually disappear and lead to the fall of the futures price $F_{t+t}^Q$ back to its equilibrium level $F_{t+t}$. The technical frictions, physical constraints and information asymmetry in cross market transactions that initially impede traders to immediately arbitrage will diminish across time, normally within a very short period. As a result, once the information of arbitrage opportunity becomes perfectly efficient, the arbitrage opportunity will no longer persist (Fama, 1970; Hull, 2012). Employing the basic econometric concepts introduced by Muth (1961) and Gujarati and Porter (2009), the term $\delta_t$ can be generalised in the form as follows:

$$\delta_t = \sum_{k=1}^{n} w_k \varepsilon_k$$  \hspace{1cm} (12)

where $\varepsilon_k \sim N(0, \sigma_k^2)$, $E(\varepsilon_i) = 0$ and $E(\varepsilon_i \varepsilon_j) = 0$ if $i = j$, $E(\varepsilon_i \varepsilon_j) = 0$ if $i \neq j$.

$\delta_t$ is composed of $n$ number of factors that are normally and independently distributed random variables $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_k\}$ that account for numerous destabilising financial shocks in the futures market. As the deviation of the futures price from its equilibrium is temporary as explained by the no-arbitrage equilibrium, the expected value of $\delta_t$ is zero. In equilibrium state, the futures price equals its fair value illustrated in equation (13).

$$E\left(F_{t+t}^Q\right) = F_{t+t}$$  \hspace{1cm} (13)

In the theoretical model, the deviation of the futures price from its equilibrium as a result of destabilising financial shock is supposed to exist in a very limited period but soon will turn out to be zero owing to the efficient market hypothesis that eventually deters away any opportunity of arbitrage.
The dynamics of crude oil price movements

Integrating the price dynamics of the spot and the futures markets, the cross market equilibrium should fulfill the following condition, as represented by equations (14) and (15), where the opportunity of arbitrage is strictly discarded (Hull, 2012):

\[ P_t e^{rc_t} = F_{t+} e^{\omega} \leftrightarrow P_t e^{rc_t} - F_{t+} e^{\omega} = 0 \]  \hspace{1cm} (14)

\[ P_t = F_{t+} e^{rc} \text{ where } \varphi = r + c \]  \hspace{1cm} (15)

where \( r \) represents the interest rate. \( c \) is the parameter that measures the cost of holding physical crude oil, such as storage fees, insurance, transport etc., and in the model, is expressed as a constant proportion of the spot price in percentage. \( \omega \) represents the convenience yield that is also written in the form of a constant proportion of the futures price. The sum of interest rate and the cost of storage is denoted by \( \varphi \) that is also known as the ‘cost of carry’. Equations (14) and (15) are exponential functions since the interest rate, the costs of storage and the convenience yield are assumed to be continuously compounded.

Substituting equation (15) into equation (9), the model equation of the dynamic price expectation change is represented by equation (16). The change of price expectation exhibits a positive correlation with the change of the futures price. As the futures price is taken as a reliable reference price, a turbulent financial market that leads to a rapidly volatile futures price will inevitably transmit the price pressure to the spot market through the price expectation channel.

\[ \frac{\Delta P_t}{P_t} = \frac{b_1 e^{rc_t} + \frac{e^{rc_t} a}{\theta} (F_t)^{-1} + b_2 e^{\omega} \frac{F_{t+}}{\theta} + e^{\omega} \frac{F_1}{\theta} \times F_t}{F_t} \]  \hspace{1cm} (16)

Equation (16) can be re-arranged in the following form, such that:

\[ \frac{\Delta P_t}{P_t} = \Phi_s + \Theta_s (F_t)^{-1} + \Phi_d \frac{F_{t+}}{\theta} + \Omega \frac{\Delta F_t}{F_t} \]  \hspace{1cm} (17)

where \( \Phi_s = b_1 e^{rc_t}, \Theta_s = \frac{e^{rc_t} a}{\theta}, \Phi_d = \frac{b_2 e^{\omega}}{\theta}, \Omega = \frac{e^{\omega}}{\theta} \)

Price expectation, equation (17), exhibits a positive correlation to a futures price movement. In retrospect to the assumption made earlier, the futures price, being a signalling indicator, tends to move in advance in response to any change in economic and financial market fundamentals. As illustrated by equations (2) and (16), since the dynamics of the spot price is predicted by the frequent variation of price expectation, it is also true that the spot price movement is subject to the futures price variation from time to time. In addition, the effect of a rapid change in price expectation arisen from a volatile futures price is extended to influence the economic fundamental changes. The interaction of all these factors is to be justified in the following sections.

To critically examine the linear dependencies among time series variables, namely, spot price \( (P) \), futures price \( (F) \), quantity supply \( (Q) \), world gross output or GDP \( (Y) \) and inventory demand \( (I) \), a Vector Error Correction Model (VECM) with the following representation as stated in equation (18) is adopted.

\[ X_t = C + GE_t + HE_{t-1} + RX_{t-1} + SX_{t-2} + VX_{t-3} + U_t \]  \hspace{1cm} (18)
$X_t$ is a vector that consists of the above-mentioned endogenous variables which are integrated of order one or in the first difference form at period $t$. In short, it is known as the vector of dependent variables in the system. Meanwhile, $X_{t-1}, X_{t-2}$ and $X_{t-3}$ are the vectors of these variables’ lags, say, at period $t - 1$, period $t - 2$ and period $t - 3$, respectively. $E_1$ and $E_2$ are the vectors which contain co-integrating equations that explain the equilibrium relationship among some of these variables in long-run. The co-integrating equations show some linear combinations of the variables which are integrated of order zero. $C$ is the vector of constant terms which represents the intercept coefficients estimates. $R$, $S$ and $V$ are $5 \times 5$ matrices which comprise the estimates of slope coefficients corresponding to lags in $X_{t-1}, X_{t-2}$ and $X_{t-3}$. On the other hand, $G$ and $H$ are $5 \times 1$ matrices of coefficient estimates that measure the gravitations of the model or $X_t$ towards its long-run equilibrium. $U_t$ represents the vector of error terms corresponding to each single regression equation and reflects the random shocks received by the model’s system. The above generalised unrestricted VAR model equation can be expanded in the following matrix form such that:

$$
\begin{bmatrix}
\Delta L_P \\
\Delta L_F \\
\Delta L_{Qd} \\
\Delta L_{Qs} \\
\Delta L_t
\end{bmatrix} =
\begin{bmatrix}
c_1 \\
c_2 \\
c_3 \\
c_4 \\
c_5
\end{bmatrix} +
\begin{bmatrix}
g_1 \\
g_2 \\
g_3 \\
g_4 \\
g_5
\end{bmatrix} \left[ L_{P,t-1} - \alpha_1 L_{F,t-1} - \alpha_2 L_{Y,t-1} - \alpha_3 L_{Q_{d,t-1}} - \alpha_4 L_{Q_{s,t-1}} \right] +
\begin{bmatrix}
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4 \\
\beta_5
\end{bmatrix} \left[ L_{F,t-1} - \beta_1 L_{Y,t-1} - \beta_2 L_{L,t-1} \right] +
\begin{bmatrix}
s_1 \\
s_2 \\
s_3 \\
s_4 \\
s_5
\end{bmatrix} \Delta L_{P,t-2} +
\begin{bmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4 \\
v_5
\end{bmatrix} \Delta L_{F,t-2} +
\begin{bmatrix}
\Delta L_{Q_{d,t-2}} \\
\Delta L_{Q_{s,t-2}} \\
\Delta L_{Y,t-2} \\
\Delta L_{L,t-2}
\end{bmatrix} +
\begin{bmatrix}
\mu_1 \\
\mu_2 \\
\mu_3 \\
\mu_4 \\
\mu_5
\end{bmatrix}
$$

(19)

All the variables in equation (19) are logarithms on quarterly basis. The period in review starts from 1990 Q3 to 2013 Q3 with 94 number of observations. FOB Brent crude oil price (USD/barrel) instead of WTI Crude Price Index is adopted for $L_P$, since the former is much more popularly used for price referencing than the latter (Lombardi and Robays, 2011). The ICE Brent Crude Continuous three month-ahead futures price (USD/barrel) is used for $L_F$. Both $L_{Q_{d,t}}$ and $L_{Q_{s,t}}$ stand for the world crude oil demand and production,
The dynamics of crude oil price movements

respectively, that are in terms of million tons of oil equivalent. \( L_Y \) represents the world gross output. It is interesting to note that \( L_Y \) is technically used as a proxy for world quantity demanded of crude oil \( LQ_d \) in order to avoid perfect multi-collinearity problem. \( L_I \) (million barrels) represents the US crude oil stockpile that has been conventionally used as one of the primary referencing indicators that influence the market participants’ investment decision making.

Data for the Brent Crude’s spot price, \( L P_s \), is obtained from the US Energy Information Administration (EIA) database. The data for ICE Brent Crude Continuous three month-ahead futures price, \( L F_s \), is collected from Open Financial Data Project (OFDP) via Quandl. Both the world crude oil production, \( LQ_s \), and the world crude oil demand, \( LQ_d \), are provided by Oxford Economics and originated from EIA database. \( L_Y \) is sourced from the Oxford Economics database. The data for \( L_I \) is provided by Thomson Reuters’ Datastream while its source is the US Department of Energy.

The raw time series data of these variables are proven non-stationary according to the results presented by Augmented Dickey-Fuller (ADF) Test, Phillips-Perron (PP) Unit Root Test and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) Unit Root Test. To rectify such statistical weakness, the time series variables turn out to be integrated of order one, as shown in equations (18) and (19). Apart from that, employing the standards of Schwarz Information Criterion and Hannan-Quinn Information Criterion, the optimal lag length selected for the regression run is three.

To construct the co-integrating model equations of prices, as justified by Pesaran and Shin (1999) who argue that the number of restrictions imposed on the model exactly equals the squared number of co-integrating equations, four restrictions are imposed based on the prior economic assumptions. Spot price, \( L P_s \), is normalised at the left-hand side of the first co-integrating equation with the additional restriction imposed such that the coefficient estimate of inventory, \( L I \), is equal to zero. It makes common sense that inventory is a pillar determinant of total quantity demanded and thus price, as shown in equation (7). However, such a concept should be carefully interpreted while modelling the co-integrating equation. Imposing such a restriction is reasonable as the inventory is a stock variable that does not show a direct statistically significant correlation to the market price movement. Instead, it is the change of inventory, \((L I - L Y_{t-1})\) itself explains the variation of market price, \((L P_s)\). In a parallel sense, the change of spot price \(\left(\frac{\Delta P}{P}\right)\) responds to the rate of change of inventory \(\left(\frac{\partial^2 I}{\partial t^2}\right)\). The second co-integrating equation represents the futures price, \(L F_s\), with another two additional restrictions imposed, such as normalisation of \(L F_s\) and statistically insignificant \(L P_s\) as the theoretical model assumes that the effect of \(L F_s\) on \(L P_s\) is unilateral in the long-run.

Last but not least, Simple Bivariate Wiener-Granger Causality Test and VAR Granger Causality/Block Exogeneity Wald Tests are employed to examine the causality among time series variables.

4 Results and discussions

The Johansen Co-integration Test affirms that there are two co-integrating equations that explain the interactions among those time series variables in the model. In line with the
economic rationales under the assumption of the theoretical model, certain restrictions are crafted to model the dynamic spot and futures prices movements of crude oil in the long-run. Pesaran and Shin (1999) suggest that the number of restrictions imposed on the model exactly equals the squared number of co-integrating equations. As the Johansen Test finds that the number of co-integrating equations is two, there are four restrictions in the model.

Constructing the first co-integrating equation for modelling spot price $LP_t$, two restrictions are imposed, such that $LP_t$ is the normalised and coefficient estimate of inventory stock $L_t$ equals zero, implying that it has no statistically significant impact on $LP_t$. As explained earlier, the change of spot price, $\left( LP_t \text{ or } \frac{\Delta P_t}{P} \right)$, is attributed to the rate of change of inventory stockpile $\left( L_t - LI_{t-1} \text{ or } \frac{\partial L_t}{\partial L_t} \right)$ rather than the change of inventory $\left( LI_t \text{ or } \frac{\partial I_t}{I_t} \right)$. Such reasoning is further supported by a causality tests in Tables 1 and 2. $LP_t$ is normalised at the left-hand side and finally the study comes up with the following long-run spot price model equation, such that:

$$LP_t = 0.7075 + 0.9516^{\text{LF}} + 1.6524^{\text{LY}} - 0.867^{\text{LQs}} - 0.0115t$$

$$[\text{Standard error}] [0.0358] [0.8890] [0.3928] [0.0069]$$

$$\text{(T-statistics)} (26.5810)^{*} (1.8587)^{**} (-2.2072)^{*} (-1.6670)^{**}$$

Table 1 The results of simple bivariate Wiener-Granger causality test

<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>df (n-k)</th>
<th>F-statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LY$ does not Granger-cause $LF$</td>
<td>90</td>
<td>10.118</td>
<td>0.000*</td>
</tr>
<tr>
<td>$LF$ does not Granger-cause $LY$</td>
<td>90</td>
<td>5.160</td>
<td>0.003*</td>
</tr>
<tr>
<td>$LI$ does not Granger-cause $LF$</td>
<td>90</td>
<td>0.875</td>
<td>0.458</td>
</tr>
<tr>
<td>$LF$ does not Granger-cause $LI$</td>
<td>90</td>
<td>1.836</td>
<td>0.147</td>
</tr>
<tr>
<td>$LP$ does not Granger-cause $LF$</td>
<td>90</td>
<td>0.508</td>
<td>0.678</td>
</tr>
<tr>
<td>$LF$ does not Granger-cause $LP$</td>
<td>90</td>
<td>2.174</td>
<td>0.097**</td>
</tr>
<tr>
<td>$LQs$ does not Granger-cause $LF$</td>
<td>90</td>
<td>2.084</td>
<td>0.109</td>
</tr>
<tr>
<td>$LP$ does not Granger-cause $LQs$</td>
<td>90</td>
<td>4.860</td>
<td>0.004**</td>
</tr>
<tr>
<td>$LI$ does not Granger-cause $LY$</td>
<td>90</td>
<td>0.880</td>
<td>0.455</td>
</tr>
<tr>
<td>$LY$ does not Granger-cause $LI$</td>
<td>90</td>
<td>3.099</td>
<td>0.031*</td>
</tr>
<tr>
<td>$LP$ does not Granger-cause $LY$</td>
<td>90</td>
<td>5.396</td>
<td>0.002*</td>
</tr>
<tr>
<td>$LY$ does not Granger-cause $LP$</td>
<td>90</td>
<td>10.948</td>
<td>0.000*</td>
</tr>
<tr>
<td>$LQs$ does not Granger-cause $LY$</td>
<td>90</td>
<td>1.512</td>
<td>0.217</td>
</tr>
<tr>
<td>$LY$ does not Granger-cause $LQs$</td>
<td>90</td>
<td>7.802</td>
<td>0.000*</td>
</tr>
</tbody>
</table>
Table 1  The results of simple bivariate Wiener-Granger causality test (continued)

<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>df (n-k)</th>
<th>F-statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP does not Granger-cause LI</td>
<td>90</td>
<td>1.753</td>
<td>0.163</td>
</tr>
<tr>
<td>LI does not Granger-cause LP</td>
<td>90</td>
<td>1.159</td>
<td>0.331</td>
</tr>
<tr>
<td>LQS does not Granger-cause LI</td>
<td>90</td>
<td>3.788</td>
<td>0.013*</td>
</tr>
<tr>
<td>LI does not Granger-cause LQS</td>
<td>90</td>
<td>0.089</td>
<td>0.966</td>
</tr>
<tr>
<td>LQS does not Granger-cause LP</td>
<td>90</td>
<td>1.849</td>
<td>0.145</td>
</tr>
<tr>
<td>LP does not Granger-cause LQS</td>
<td>90</td>
<td>4.465</td>
<td>0.006*</td>
</tr>
</tbody>
</table>

Notes: *Rejecting null hypothesis at the 5% significance level.
**Rejecting null hypothesis at the 10% significance level.
Null Hypothesis: independent variable does not Granger-cause dependent variable. Number of lags is three. All the figures are rounded up to the nearest three decimal places.

Table 2  The results of VAR Granger causality/block exogeneity Wald test

<table>
<thead>
<tr>
<th>VAR Granger causality/block exogeneity Wald tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample period: 1990 Q3–2013 Q3</td>
</tr>
<tr>
<td>Number of observations: 94</td>
</tr>
</tbody>
</table>

<p>| Dependent variable: LP_t                      |</p>
<table>
<thead>
<tr>
<th>Excluded</th>
<th>Chi-sq</th>
<th>Df</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF_t</td>
<td>6.268</td>
<td>3</td>
<td>0.098**</td>
</tr>
<tr>
<td>LY_t</td>
<td>27.370</td>
<td>3</td>
<td>0.000*</td>
</tr>
<tr>
<td>LI_t</td>
<td>2.865</td>
<td>3</td>
<td>0.413</td>
</tr>
<tr>
<td>LQs,t</td>
<td>4.833</td>
<td>3</td>
<td>0.184</td>
</tr>
<tr>
<td>All</td>
<td>37.064</td>
<td>12</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

<p>| Dependent variable: LF_t                      |</p>
<table>
<thead>
<tr>
<th>Excluded</th>
<th>Chi-sq</th>
<th>Df</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LY_t</td>
<td>21.258</td>
<td>3</td>
<td>0.00*</td>
</tr>
<tr>
<td>LI_t</td>
<td>2.905</td>
<td>3</td>
<td>0.407</td>
</tr>
<tr>
<td>LP_t</td>
<td>0.667</td>
<td>3</td>
<td>0.881</td>
</tr>
<tr>
<td>LQs,t</td>
<td>2.010</td>
<td>3</td>
<td>0.570</td>
</tr>
<tr>
<td>All</td>
<td>26.303</td>
<td>12</td>
<td>0.010*</td>
</tr>
</tbody>
</table>

<p>| Dependent variable: LY_t                      |</p>
<table>
<thead>
<tr>
<th>Excluded</th>
<th>Chi-sq</th>
<th>Df</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF_t</td>
<td>0.427</td>
<td>3</td>
<td>0.935</td>
</tr>
<tr>
<td>LI_t</td>
<td>2.343</td>
<td>3</td>
<td>0.504</td>
</tr>
<tr>
<td>LP_t</td>
<td>1.397</td>
<td>3</td>
<td>0.706</td>
</tr>
<tr>
<td>LQs,t</td>
<td>0.856</td>
<td>3</td>
<td>0.836</td>
</tr>
<tr>
<td>All</td>
<td>14.549</td>
<td>12</td>
<td>0.267</td>
</tr>
</tbody>
</table>
Table 2  The results of VAR Granger causality/block exogeneity Wald test (continued)

<table>
<thead>
<tr>
<th>Dependent variable: ( L_i )</th>
<th>Excluded</th>
<th>Chi-sq</th>
<th>Df</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( LF_t )</td>
<td>2.888</td>
<td>3</td>
<td>0.409</td>
<td></td>
</tr>
<tr>
<td>( LY_t )</td>
<td>3.342</td>
<td>3</td>
<td>0.342</td>
<td></td>
</tr>
<tr>
<td>( LP_t )</td>
<td>1.212</td>
<td>3</td>
<td>0.750</td>
<td></td>
</tr>
<tr>
<td>( LQ_{s,t} )</td>
<td>0.765</td>
<td>3</td>
<td>0.858</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>12.724</td>
<td>12</td>
<td>0.389</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent variable: ( LQ_{s,t} )</th>
<th>Excluded</th>
<th>Chi-sq</th>
<th>Df</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( LF_t )</td>
<td>13.390</td>
<td>3</td>
<td>0.004*</td>
<td></td>
</tr>
<tr>
<td>( LY_t )</td>
<td>10.578</td>
<td>3</td>
<td>0.014*</td>
<td></td>
</tr>
<tr>
<td>( LQ_{s,t} )</td>
<td>0.203</td>
<td>3</td>
<td>0.977</td>
<td></td>
</tr>
<tr>
<td>( LP_t )</td>
<td>11.558</td>
<td>3</td>
<td>0.009*</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>35.295</td>
<td>12</td>
<td>0.000*</td>
<td></td>
</tr>
</tbody>
</table>

Note:  *Rejecting null hypothesis at the 5% significance level.
**Rejecting null hypothesis at level of significance 10%.
Null hypothesis: independent variable does not Granger-cause dependent variable. Number of lags is three. All the figures are rounded up to nearest three decimal places.

In the long-run, the time series of \( LP_t \) is said to be a random walk model (non-stationary) with drift and deterministic trend. The time factor \( t \) is measured chronologically. The above long-run spot price model represented by co-integrating equation (20) which is derived from the VECM indicates that the coefficients estimates of \( LF_t \) and \( LQ_{s,t} \) are significantly different from zero at 5% level of significance while the coefficients estimates of \( LY_t \) and time factor \( t \) are significantly different from zero at 10% level of significance. The results are in line with the study’s expectation that the spot price movement is statistically correlated to the explanatory variables. \( LF_t \) has a positive sign, implying that any 1% increase (or decrease) in the futures price will transmit the upward (or downward) price pressure of 0.9516% on \( LP_t \). \( LY_t \) is positively correlated to \( LP_t \), indicating that each 1% increase (or decrease) in aggregate demand owing to flourishing economic activities will boost (or reduce) the quantity demanded for crude oil, ceteris paribus and hence the spot price goes up (or down) by 1.6524%. The correlation between \( LQ_{s,t} \) and \( LP_t \) is negative as each 1% increase (or decrease) of production supply, holding other variables constant, leads to a fall (or rise) of the spot price by 0.867%. Moreover, it is interesting to note that the time series data exhibits a negative time trend. There are two implications to justify such a finding. First, the first co-integrating model equation’s error terms that represent the deviations from the trend line are purely random and thus dies out quickly across the time periods (Gujarati and Porter, 2009). Second, the magnitude of the change of spot price \( LP_t \) is diminishing, say, by approximately 0.01% as the time goes by. Such a finding may be attributed to the growing consumption of crude oil substitutes as a result of the proliferation of clean energy use and rapidly improved...
The dynamics of crude oil price movements

fuel efficiencies since 1990. All else equal, the technological progress and the rising awareness of carbon emissions reductions that cut down the market’s dependence on crude oil in some industries lead to a downside potential of crude oil price across the periods. In a nutshell, the spot price movement is driven by the market participants’ price expectation change from time to time that is predicted based on the futures price variation in the financial market. The futures price is typically taken as a reliable indicator or reference price to form market participants’ price expectations and then determines the spot price trend. Also, in the long-run, the effect of \( LF_t \) on \( LP_t \) is less influential than that of \( LY_t \).

Apart from equation (20), the second co-integrating equation is derived from VECM with another two restrictions, such that \( LF_t \) is normalised and the coefficient estimate of the spot price \( LP_t \) equals zero based on the theoretical model assumption in the preceding chapter that \( LF_t \) has a unilateral impact on \( LP_t \) to model the futures price movement in the long-run. Normalising \( LF_t \) on the left-hand side, the equation is exhibited as follows:

\[
LF_t = 12.9969 + 15.4848LY_t + 2.4423LI_t - 3.0854LQ_{s,t} - 0.0988t
\]

\[
(T – statistics) (4.2385)* (5.7223)* (1.2644) (3.0683)*
\]

where *means statistically significant at 5%; **means statistically significant at 10%.

The results from equation (21) show that the coefficients estimates of most explanatory variables, except for that of \( LQ_{s,t} \), are significantly different from zero at the level of significance 5%. Their signs comply with the expectation and make economic sense. Based on the study’s assumption, market participants, both commercial traders and financial investors, who trade crude oil futures typically, make their transaction decisions based on several macroeconomic considerations, such as global economic growth, strategic inventory stockpile hold by major countries (e.g., the USA, China, etc.) and production. Any economic shock arisen from the change of each above-mentioned factor is supposed to influence their trading positions as well as the monetary amount they would like to pour into the futures market.

For each 1% increase in \( LY_t \), \( LF_t \) will increase by 15.4848%, holding other variables constant. Such a positive correlation between \( LY_t \) and \( LF_t \) is explained by the fact that an economic boom associated with the high demand for crude oil will surely drive the price up. Furthermore, the 1% increase in inventory stockpile is followed by 2.4423% fall in \( LF_t \). In economic sense, futures traders would expect the crude oil market may not be tight in the near future and hence it is not profitable enough for them to speculate on futures trading at the current moment. As the volume of transactions cools down in the futures market, \( LF_t \) will drop accordingly. Similarly, the regression in equation (20) also exhibits a negative time trend that is the same as equation (19). Since the futures price is a function of spot price in an equilibrium state, it has a tendency to go down across the periods, holding other variables unchanged. \( LQ_{s,t} \) does not perform a statistical significant correlation to \( LF_t \).

Having rectified equation (20), the results are presented as follows:

\[
LF_t = -22.5026 + 12.5257LY_t + 2.9304LI_t - 0.0794t
\]

\[
(T – statistics) (3.0131)* (-5.7551)* (-2.0639)*
\]

where *means statistically significant at 5%; **means statistically significant at 10%.
In equation (22), it is apparent to see that all the coefficient estimates of the explanatory variables are significantly different from zero at level of significance 5% with correct signs that meet the model’s expectation. Moreover, the magnitudes of coefficients estimates slightly change, such that the slope coefficient of $LY_t$ changes from 15.4848 to 12.5257 while that of $LI_t$ changes from 2.4423 to 2.9304.

**Figure 1** Impulse response functions (see online version for colours)

![Impulse response functions](image1)

Notes: The combined graph consolidates the impulse response of spot price to the variations of changes of other variables across the time periods. Each time period represents a quarter.

**Figure 2** Variance decomposition (see online version for colours)

![Variance decomposition](image2)

Notes: The combined graph indicates how large the impulse or shock of each variable accounts for the variation of fluctuation of spot price in term of percentage across the time periods. Each time period represents a quarter.
The results of the impulse response function and the variance decomposition analyses are presented in Figures 1 and 2, respectively. In Figure 1, the magnitude of the response of the spot price to the change of the futures price is much larger than is the world gross output in the short-run, say for the first two quarters. As time goes by, the spot price movement turns out to be more sensitive to the change of the world gross output than to that of the futures price. Transmitting from the short-run to the long-run, the world gross output becomes increasingly influential to the spot price movement while the effect of the futures price movement on the variation of the spot price eventually falls and becomes relatively less significant, comparing to world gross output. Such a finding can be explained by the fact that any shock arisen from the unanticipated change of world gross output takes longer time to materialise and be reflected in the spot price trend.

On the other hand, the analysis of the variance decomposition also presents quite similar results. In the short-run, the futures price is the key factor that contributes most to the volatility of the spot price as compared to other variables. As presented in Figure 2, for the first three quarters, the shock of spot price primarily accounts for the volatility of itself. In the following quarters, futures price gradually emerges to be the dominant factor that attributes to the variation of spot price volatility. The lagged effect of futures price on spot price can be explained by the market imperfections in reality, to name a few, the presences of imperfect information accessibility and the frictions of carrying out transactions. The results of variance decomposition just exactly comply with the theoretical model’s hypothesis, as shown in equation (15), that the prediction of spot price trend is always based on the futures price movement in the equilibrium state. Therefore, the fluctuation of futures price plays the most critical role relative to other variables in influencing the spot price variation at most of the times. The financialisation of crude oil prices and the popularisation of trading of financial derivatives appear to accelerate the volatility of futures price and hence spot price.

Last but not least, the results of Simple Bivariate Wiener-Granger Causality Test (WGCT) and VAR Granger Causality/Block Exogeneity Wald Test (VARWT), as presented in tables 1 and 2, more or less yield similar results but with slight differences. To make things concise, the study only focuses on the results of several key variables that are relevant to the justification of the theoretical model. The results of WGCT indicate that, $LY_t$ and $LF_t$ as well as $LP_t$ and $LY_t$ perform a bilateral Granger causality with one another as their null hypotheses are rejected at the level of significance 5% (See Table 1). In economics sense, the change in global output level or GDP which is also a proxy of crude oil demand will result in proportional changes of both spot and futures prices, ceteris paribus. At the meantime, the investment and consumption decisions will vary accordingly in response to the prices changes. For example, an unprecedented crude oil price hike substantially increases the cost of production and thus the overall output falls; otherwise, the opposite is true. Therefore, the change of crude oil price remains dynamic at all times. Furthermore, $LF_t$ Granger-causes $LP_t$ at the 10% level of significance with the $p$-value of 0.097. This exactly conforms to the rationale of equation (15) such that futures price is a forward-looking price referencing benchmark that to certain extent predicts the spot price in the following period. $LP_t$ and $LF_t$ predict $LQ_s$ with $p$-values of 0.006 and 0.004, respectively, when the level of significance is set to be 5%. On the other hand, VARWT shows that $LY_t$ Granger-causes $LF_t$ and $LP_t$, at level of significance 5%, with the $p$-values that approach zero $9.31 \times 10^{-5}$ and
4.92 \times 10^{-5}$, but the reverse is not statistically significant. Furthermore, three variables, namely $LP_t$, $LF_t$ and $LY_t$, simultaneously predict $LY_{s,t}$ at the 5\% level of significance with the $p$-values of 0.009, 0.004 and 0.014, respectively. The Granger Causality Tests' results comply with the study’s expected outcomes, though; the study notes that the tests are conducted by using the raw data of each variable that is non-stationary.

Key notable findings from the results are as follows. First, the spot price movement is predicted by the futures price from time to time as shown in the results of Granger Causality Test. The impulse response of the spot price to the change of the futures price is even larger than to the change of GDP in the short-run. Also, the short-term spot price volatility is primarily determined by the variation of the futures price. Nevertheless, the effect of the futures price on the spot price diminishes as time passes to medium-term and long-term. The futures price is widely used by market participants as a reliable signalling indicator to predict the spot price trend in the near future.

Second, gross output remains a significant fundamental factor that determines both the futures and the spot prices movements. The impulse responses diagram indicates that the effect of the GDP on the spot and the futures prices is getting larger over time. Similar outcomes appear in the analyses of the spot price and the futures price variance decompositions as well. Any significant change of GDP may not lead to an immediate impact to the prices movements within a short period but the impact will materialise and grow larger later. Such a finding complies with the regression result of the co-integrating equation that has found that the effect of GDP on the spot price far exceeds that of the futures price in the long-run.

Third, the Granger Causality test shows that the spot price, the futures price and economic activities (typically measured by GDP) play an influential role in determining the quantity supplied of crude oil. Such a finding exactly conforms to the rationale of equation (1) in the theoretical model. Positive shocks from the spot and the futures market as well as the macroeconomic environment will raise the producers’ price expectation and hence revise their quantity supplied. This implies that the change of price expectation not only facilitates the price pressure transmission across markets, but also significantly affects the fundamental factors of the market.

The study has constructed a model to illustrate how the spot market and the futures market are inter-related. The price pressure from the former as a result of changes in economic and financial fundamentals will be quickly and effectively transmitted to the latter through a price expectation channel. As shown in Figure 3, having observed the spot and futures prices movement for the past 23 years, the volatility of prices has intensified, especially in the most recent years. It is apparent to see the gap between daily highest and lowest prices as well as the difference between daily price and the settlement price for each day in the futures market has been widening. Such an observation is coincidentally in line with an enormous rise of transaction volumes and open interests caused by the financialisation of crude oil prices. Commodities emerged to be a lucrative alternative investment asset that allowed more opportunities for portfolio diversification and better risk-return trade-off (Gorton and Rouwenhorst, 2006). An enduring bullish futures trading that started in the beginning of the 2000s plays a catalytic role in leading the market participants to expect for a more price upward potential and thus is the culprit of recently skyrocketing futures price (Ripple and Moosa; 2009). Otherwise, vice versa.
The dynamics of crude oil price movements

Reconciling the findings with the study’s theoretical model of price expectation, there is no doubt that crude oil financialisation in the 1990s that results in the increasing frequency of trading in futures market attributes to the price hike in the spot market (see Figure 4). Since market participants in the spot market tend to use futures price as a reliable price reference indicator to predict the future trend, any drastic change of futures price soon leads to rapid change of their price expectation and thus their investments decisions. As such, via the price expectation channel, the adjustments of demands and supplies will accordingly make the spot price exhibit the same volatility as does the futures price. It is not surprising to say that the financialisation of crude oil is to blame for destabilisation of both futures and spot markets. At the other intellectual end, conservative studies advocates that only fundamental factors remain relevant to explain the price variation in an equilibrium market context from time to time. If such an argument had been absolutely true, the daily price gap, a simple proxy to measure volatility, would have not exhibited increasing volatility as the time moves nearer to the end of period in review. Instead, the magnitudes of prices fluctuations should have been consistently moderate as time goes by.

Referring to Figure 5, the study has identified that the share of crude oil reserves of OPEC exceeds that of non-OPEC. Starting from 1999, non-OPEC producers’ crude oil reserves have been depleting at an increasing rate relative to OPEC producers owing to massive extraction in recent decades. Moreover, Figure 6 illustrates that OPEC’s crude oil reserves have been going up significantly while the growth of non-OPEC’s crude oil reserves has been stagnating or increasing at a negligibly low rate. Perhaps in the farther future, the fall of non-OPEC reserves will result in the loss of competition in the global market. It is expected that the OPEC will probably once again regain the monopoly power that allows it to significantly influence the global crude oil pricing. By the time,
the existing model used in the study, which is based on the competitive and efficient market assumptions, as well as the empirical results thus may not be applicable in the presence of foreseeable structural change of global crude oil market. Such possibility should be addressed in a future study about the dynamics of crude oil prices if it turns out to be true.

**Figure 4** Price co-movement for spot and future markets (see online version for colours)

![Price Co-movement for Spot and Futures Markets](image_url)

Notes: The chart is depicted by integrating the historical spot and futures price movements across periods as well as the open interest. The vertical axes at the left and right hand sites represent price and number of contracts traded daily respectively. The horizontal axis represents time period.

**Source:** Intercontinental Exchange (ICE) and International Energy Agency (IEA)

**Figure 5** World prove crude oil reserves percentage share of total (see online version for colours)

![World Prove Crude Oil Reserves Percentage Share of Total](image_url)

Notes: The diagram shows the percentage share of crude oil reserves of OPEC and Non-OPEC producers from 1980 to 2012. The OPEC producers’ share of reserves is much larger than that of non-OPEC producers. In the first decade of the 21st century, the non-OPEC share of reserves is diminishing vis-à-vis OPEC producers. For example, the non OPEC’s crude oil reserves constitute 28% of the world’s reserves in 1999. However, it has gradually dropped for the following 13 years to 22% in 2012.

**Source:** BP Statistical Review
The dynamics of crude oil price movements

5 Conclusion

This study examines the correlation of the spot price to the futures price and other fundamental factors. The results show that the growth of the spot price exhibits a positive correlation to that of the futures price with a coefficient estimate of approximately one. This complies with the actual observation that the change of the spot price responses to the change of the futures price with the equally same amount in the market. The study notes that the effect of aggregate demand on the spot price is much larger than that of the futures price in the long-run. The results from causality tests affirm that the variation of the futures price predict the spot price movement from time to time. Such a finding indicates that the dynamics of the spot and the futures prices are integrated through the price expectation channel.

Futures market is always responsive to both fundamental shocks and financial market shocks. In the absence of the changes of economic fundamental factors, speculation hype in the futures market will suffice to significantly destabilise the futures price and the upward price pressure will very soon be proportionally transmitted to the spot market through the rational price expectation channel. The change in the price expectation will further affect the market participants’ investment decisions, say, crude oil supply and hence the other market fundamental factors may change accordingly. Consequently, the spot price movement is drastically destabilised. Such a theoretical inference is in line with the actual price movement in the present market. As the volume of transactions and the amount of open interest in the futures market substantially increase as a result of crude oil financialisation, the effect of huge fluctuation of the futures price on the spot
market is increasingly significant. Thus, the crude oil price will become more volatile and more uncertain than it was in the past. The forecast of the spot price trend in the future that solely relies on a fundamental analysis may lack accuracy.

The study does foresee that the gradually deteriorating imbalance of crude oil reserves and future production share between OPEC and non-OPEC is expected to make a structural change to the existing competitive market. It takes much longer time for these possible effects to materialise but taking into account such considerations would be an interesting study.

References


Appendix

1 Derivation of equation (9)

Suppose the market is in the state of equilibrium, total supply of oil \( (Q_s^t) \) equals total demand \( (Q_d^t) \) of oil, such that:

\[
Q_s^t = Q_d^t \\
Q_{s,t} = Q_{d,t} + (I_t - I_{t-1}) \\
\alpha_s + \beta_s P_t^r + \mu_{s,t} + \lambda (P_t^r - P_t) = \alpha_s \beta_s P_t + \mu_{s,t} + \lambda (P_t^r - P_t) \\
\alpha_d + \beta_d P_t^r + \mu_{d,t} + \lambda (P_t^r - P_t) = \alpha_d \beta_d P_t + \mu_{d,t} + \lambda (P_t^r - P_t) \\
\lambda (P_t^r - P_t) = \alpha_s - \alpha_d + \beta_s P_t^r + \beta_d P_t - \lambda P_t - \lambda P + \mu_s - \mu_d \\
\Delta P_t^r = \frac{(\alpha_s - \alpha_d) + \beta_s P_t^r + \beta_d P_t - \lambda P_t - \lambda P + \mu_s - \mu_d}{\lambda}
\]

For simplicity, we assume that:

\[
a = \frac{\alpha_s - \alpha_d}{\lambda} \\
b_i = \frac{\beta_i}{\lambda} \text{ for } i \in \{s,d\} \\
u = \frac{\mu_{s,t} - \mu_{d,t}}{\lambda}
\]

Therefore, we get:

\[
\Delta P_t^r = a + b_s P_t^r + b_d P_t + \Delta P_{t-1} + u
\]

where \( u \sim N(0, \sigma_u^2) \).

The change in expected prices at the point of time \((t)\) is attributed to the change in actual prices at the point of time \((t-1)\), the actual price \((P_t)\), the expected price \((P_t^r)\) and others exogenous factors that are represented by \((U)\).
2 Derivation of equations (16) and (17)

Let’s see the equations of expected price change and spot price:

\[
\Delta P_t^e = A + B_t P_t + B_s P_s + \Delta P_{t-1} + U
\]
\[
P_t = F_{t,i} e^{\omega \theta}
\]
\[
\Delta P_{t-1} = P_t - P_{t-1}
\]
\[
\Delta P_t = F_{t,i} e^{\omega \theta} - F_t e^{\omega \theta}
\]

Substituting \((F_{t,i} e^{\omega \theta} - F_t e^{\omega \theta})\) for \((\Delta P_{t-1})\), we get:

\[
\Delta P_t^e = a + b_i P_t^e + b_F t e^{\omega \theta} + F_{t,i} e^{\omega \theta} - F_t e^{\omega \theta} + u
\]
\[
\Delta P_t^f = a + b_i P_t^f + b_F t e^{\omega \theta} + e^{\omega \theta} \Delta F_t^e + u
\]

Recall that:

\[
P_t^e = \theta P_{t-1}
\]
\[
P_t = F_{t,i} e^{\omega \theta}
\]
\[
F_{t,i} = P_t e^{\omega \theta}
\]
\[
\Delta P_t^f = a + b_i \theta F_t e^{\omega \theta} + b_F t e^{\omega \theta} - e^{\omega \theta} \Delta F_t^e + u
\]

Let the above equation divided by \(F_t\):

\[
\frac{\Delta P_t^e}{F_t} = a + b_i \theta F_t e^{\omega \theta} + b_F t e^{\omega \theta} + e^{\omega \theta} \Delta F_t^e + u
\]
\[
\frac{\Delta P_t^f}{e^{\omega \theta} P_{t-1}} = a + b_i \theta F_t e^{\omega \theta} + b_F t e^{\omega \theta} + e^{\omega \theta} \Delta F_t^e + u
\]
\[
\frac{\Delta P_t^f}{e^{\omega \theta} \frac{P_{t-1}}{F_t}} = a + b_i \theta F_t e^{\omega \theta} + b_F t e^{\omega \theta} + e^{\omega \theta} \Delta F_t^e + u
\]
\[
\frac{\Delta P_t^f}{P_t^e} = e^{\omega \theta} a + b_i \theta F_t e^{\omega \theta} + b_F t e^{\omega \theta} + e^{\omega \theta} \Delta F_t^e + u
\]
\[
\frac{\Delta P_t^f}{P_t^f} = e^{\omega \theta} a + b_i e^{\omega \theta} + b_F t e^{\omega \theta} + e^{\omega \theta} \Delta F_t^e + u
\]
\[
\frac{\Delta P_t^f}{P_t^e} = \frac{e^{\omega \theta} a}{\theta F_t} + b_i \frac{e^{\omega \theta}}{\theta F_t} + \frac{e^{\omega \theta} \Delta F_t^e}{\theta F_t} + b_F t e^{\omega \theta} + \frac{e^{\omega \theta} \Delta F_t^e}{\theta F_t}
\]
\[
\frac{\Delta P_t^f}{P_t^f} = \frac{e^{\omega \theta} a}{\theta F_t} + b_i \frac{e^{\omega \theta}}{\theta F_t} + \frac{e^{\omega \theta} \Delta F_t^e}{\theta F_t} + \frac{b_F t e^{\omega \theta}}{\theta F_t} + \frac{e^{\omega \theta} \Delta F_t^e}{\theta F_t}
\]

Since \(u \sim N(0, \sigma^2)\), we assume \(u \equiv E(U) = 0\), then we get:

\[
\frac{\Delta P_t^f}{P_t^e} = b e^{\omega \theta}  + \frac{e^{\omega \theta}}{\theta (F_t)^{-1}} + \frac{b_F e^{\omega \theta}}{\theta F_t} + \frac{e^{\omega \theta} \Delta F_t^e}{\theta F_t}
\]

\[
\frac{\Delta P_t^f}{P_t^f} = b e^{\omega \theta} + \frac{e^{\omega \theta}}{\theta} + \frac{b_F e^{\omega \theta}}{\theta F_t} + \frac{e^{\omega \theta} \Delta F_t^e}{\theta F_t}
\]

\[
\frac{\Delta P_t^f}{P_t^e} = b e^{\omega \theta} + \frac{e^{\omega \theta}}{\theta (F_t)^{-1}} + \frac{b_F e^{\omega \theta}}{\theta F_t} + \frac{e^{\omega \theta} \Delta F_t^e}{\theta F_t}
\]
Simplifying the equation:

\[ \Delta P'_t = a + b \theta F_t e^{\sigma \omega} + b_2 F_{t+1} e^{\sigma \omega} + e^{\sigma \omega} \Delta F_t \]

\[ \frac{\Delta P'_t}{P'_t} = \Phi_s + \Theta_s (F_t)^{-1} + \Phi_d \frac{F_{t+1}}{F_t} + \Omega \frac{\Delta F_t}{F_t} \]

where \( \Phi_s = b e^{\sigma \omega} \), \( \Phi_d = \frac{b_2 e^{\sigma \omega}}{\theta} \), \( \Theta_s = \frac{e^{\sigma \omega} a}{\theta} \), \( \Omega = e^{\sigma \omega} \).

To sum, the mechanism that the change in expected price of oil from point of time \( t \) to \( (t + 1) \) can be predicted and explained by its associated change in the futures price. The model equation is in line with our hypothesis that the market participants, both commercial and non-commercial, tend to take the futures price as a reference price to form their expectation about the price movement and hence make their investment decisions rationally. As a result, the fundamentals factors, i.e. demand and supply, will adjust accordingly and a new spot price in the physical market will be formed. Hence, both the futures and the spot markets remain equilibrium.

\[ \ln P_t = \alpha + \beta (\Delta \ln I_t) \]

\[ \frac{\partial}{\partial I_t} \ln P_t = \frac{\partial}{\partial I_t} \left[ \alpha + \beta (\Delta \ln I_t) \right] \]

\[ \frac{1}{P_t} \frac{\partial P_t}{\partial I_t} = \beta \frac{\partial}{\partial I_t} \ln I_t \]

\[ \frac{1}{P_t} \frac{\partial P_t}{\partial I_t} = \beta \frac{\partial}{\partial I_t} \frac{1}{I_t} \]

\[ \frac{\partial P_t}{P_t} = \beta \frac{\partial^2 I_t}{\partial I_t^2} \]
Oil demand forecasting for India using artificial neural network

S. Jebaraj*

Department of Mechanical Engineering,
Universiti Teknologi PETRONAS,
Perak, Malaysia
Email: sjabaraj74@gmail.com
*Corresponding author

S. Iniyan

Department of Mechanical Engineering,
Anna University,
Chennai, Tamil Nadu, India
Email: iniyan777@hotmail.com

Abstract: Energy is a vital input for the growth of any nation. Since oil resource has become a vital factor for future developments of a country, a system of models has to be developed to provide forecasts of oil demands in various sectors. This analysis utilises regression techniques, double moving average method, double exponential smoothing method, triple exponential smoothing method, Autoregressive Integrated Moving Average (ARIMA) model and Artificial Neural Network (ANN) model (univariate and multivariate) for oil demand forecasts in India. Model validation is done to select the best forecasting model. It is found that the ANN model gives better results in most of the cases. Hence, it is suggested that the ANN model can be used for forecasting oil demands in India. It is also predicted that the total oil demand for the years 2020 and 2030 will be 415,373 and 720,688 thousand tonnes, respectively.

Keywords: oil consumption; demand forecasting; forecasting models; ANN; artificial neural network; model simulation.


Biographical notes: S. Jebaraj is currently working in the Mechanical Engineering Department, Universiti Teknologi PETRONAS, Malaysia. He started his teaching profession in 1996. He obtained his Doctorate degree in Mechanical Engineering in 2007 from Anna University, India. He has published ten research papers in reputed international journals and has presented more than 35 papers in conferences in the area of energy modelling. His papers received more than 725 citations. He has more than 17 years of experience in teaching in various universities. His area of research is energy planning, policy, energy modelling, renewable energy and heat transfer.
1 Introduction

Over the past two decades, forecasting has gained a widespread acceptance as an integral part of business planning and decision-making. Recent literature on energy forecasting offers a broad range of forecasting tools. Oil consumption is characterised by its large variations over a period of time. It varies widely from year to year and exhibits seasonal variations. Formulating a forecasting model that can accurately forecast oil consumption is of prime importance in energy system planning. The estimated reserves of crude oil in India as on 31 March 2012 stood at 759.59 Million Tonnes (MT). The geographical distribution of crude oil in India indicates that the maximum reserves are in the western offshore (44.46%), followed by Assam (22.71%). As on 31 March 2012, there were in total 20 refineries in the country: 17 in the public sector and three in the private sector. Public sector refineries are located in Guwahati, Barauni, Koyali, Haldia, Mathura, Digboi, Panipat, Vishakhapatnam, Chennai, Nagapattinam, Kochi, Bongaigaon, Numaligarh, Mangalore, Tatipaka and Mumbai (which has two refineries). Private sector refineries were built by Reliance Petroleum Ltd. and Essar Oil in Jamnagar and Vadinar, respectively. Total installed crude oil refining capacity in the country at the end of March 2012 was around 198 million tonnes per annum. There was an increase of 5.75% over the previous year to the installed refining capacity. Production of crude petroleum increased from 6.82 MTs during 1970–1971 to 38.09 MTs during 2011–2012. In 2011–2012, the production of petroleum products in the country was 196.71 MTs as against 190.32 MTs during 2010–2011, an increase of 3.36%. High-speed diesel oil accounted for the maximum share (41.63%), followed by motor gasoline (13.67%), fuel oil (9.89%), naphtha (8.73%), kerosene (3.8%) and aviation turbine fuel (5.11%). This paper aims at developing forecasting models, using different techniques, for oil demand in various sectors in India for the years 2020 and 2030.

A detailed literature survey has been conducted for various energy models such as energy planning models, energy supply–demand models, forecasting models and energy models based on Artificial Neural Network (ANN), the findings of which are presented here. An electricity demand forecasting model has been formulated by Zhou et al. (2006) using a trigonometric grey prediction approach. The benefits and challenges of the demand side management have been studied by Strbac (2008). A forecasting model to predict the energy demand in the transportation sector in China has been formulated by Zhang et al. (2009). A long-term energy consumption prediction model using ANNs has been developed by Ekonomou (2010). The potential of demand side management in energy-intensive industries for electricity markets in Germany has been formulated by Paulus and Borggreve (2011). Movagharnejad et al. (2011) developed a forecasting
model to evaluate the differences between various commercial oil prices in the Persian Gulf region by neural network. Electric load forecasting is carried out by Hong (2011) using seasonal recurrent Support Vector Regression (SVR) with the chaotic artificial bee colony algorithm. Future oil demand in Iran has been predicted by Behrang et al. (2011) using the Gravitational Search Algorithm (GSA). Badurally Adam et al. (2011) developed a forecasting model using the non-homogeneous Gompertz diffusion technique to determine the peak electricity demand in Mauritius. Wang et al. (2012) developed a regional electricity demand forecasting model using decomposition and statistical analysis. A short-term electric load forecasting model has been developed by Deihimi and Showkati (2012) using echo state networks. Yu and Zhu (2012) adopted a hybrid procedure for energy demand forecasting in China. Barton et al. (2013) have analysed the evolution of electricity demand and the role for demand side participation, in buildings and transport. A review of the costs and benefits of demand response for electricity in the UK has been discussed by Bradley et al. (2013). The wind energy gains, financial savings and peak time load reduction in the demand side management has been analysed by Finn et al. (2013). An et al. (2013) formulated an electricity demand forecasting model using multi-output Feedforward Neural Network (FFNN) with empirical mode decomposition-based signal filtering technique. A case study has been carried out by Zahedi et al. (2013) to estimate the electricity demand using an adaptive neuro-fuzzy network for the Ontario province in Canada. A model has been developed by Deihimi et al. (2013) to predict the short-term electric load and temperature using wavelet echo state networks with a neural reconstruction method. Gils (2014) has studied the assessment of the theoretical demand response potential in Europe. The categorisation of residential electricity consumption as a basis for the assessment of the impacts of demand response actions has been done by Soares et al. (2014). A forecasting model has been developed by Bennett et al. (2014) to predict the low voltage distribution network demand profiles using pattern recognition-based expert system. Zeng et al. (2014) has predicted the energy consumption of multiproduct pipeline using the ANN method.

It has been identified that the ANN can be used in forecasting the oil demand in the country. Even though forecasting models are already available, there are limitations of the past consumption data, Gross National Product (GNP) and population not being used collectively for India in the oil demand forecasting models. Hence, in the present research work, an attempt has been made to develop the oil demand forecasting models in various sectors by considering the population as well as GNP as input variables in addition to past oil consumption data.

2 Methodology

In the present work, an attempt was made to develop forecasting models for oil consumption in various sectors including total oil consumption in India. This analysis utilises regression techniques, double moving average method, double exponential smoothing method, triple exponential smoothing method, Autoregressive Integrated Moving Average (ARIMA) model and ANN model (univariate and multivariate) for the oil demand forecasting purpose. Figure 1 illustrates the schematic representation of the development of forecasting models. The methodologies of the above-mentioned models are discussed in the following sections.
2.1 Time series regression methods

Regression is a statistical method of fitting a line through data to minimise squared error. In the time series technique, past consumption data were used to predict future oil demands. Time series regression techniques such as a linear model, an exponential model, a power model and a quadratic model were used in the analysis. The methodologies of these models are discussed in the following sections.

2.1.1 Linear model

In the linear model, the relationship used to fit the ‘n’ number of data is a linear equation in one variable. The linear model is of the form

\[ y = ax + b \]  

where \( y \) is oil consumption in a specific sector in a particular time period, \( x \) is the time period in a year and \( a \) and \( b \) are constants.

2.1.2 Exponential model

An exponential growth curve has a constant rate of growth at any point in time, resulting in staggering increases over longer periods as the increase is compounded exponentially. This model can be approximated by the exponential curve of the following form:

\[ Y_t = e^{(a+bt)} \]
where $Y_t$ is the time series at time $t$, $e = 2.71828$, $a$ is the intercept and $b$ is the equivalent of the slope and denotes the growth rate. The above equation can be transformed to make it linear, by taking natural logarithms of both sides of the equation. The forecasting is obtained by the following equation:

$$F_{t+m} = e^{(a+bt)}$$

where $m$ is the number of periods ahead to be forecast.

### 2.1.3 Power model

The power model is also one of the time series regression techniques considered for forecasting. The power model is of the form

$$y = ax^b$$

where $y$ is oil consumption in a specific sector in a particular time period, $x$ is the time period in a year and $a$ and $b$ are constants.

### 2.1.4 Quadratic model

If the data appear to follow a quadratic pattern, then the non-linear regression has to be applied. The quadratic model is of the form

$$y = ax^2 + bx + c$$

where $y$ is the oil consumption in a specific sector in a particular time period, $x$ is the time period in a year and $a$, $b$ and $c$ are constants.

### 2.2 Double moving average method

The moving averages are applicable to time series data and in many situations they are more appropriate and easier to use than regression methods. In the present work, the double moving average method was used as one of the methods for forecasting. To calculate the double moving average, $M_{i[1]}$, the simple moving average $M_{i[0]}$ was treated as an individual data point and a moving average of these averages was obtained. The values of $M_{i[1]}$ were used to determine the equation for forecasting future oil demands. The forecasting equation is of the form

$$y_{i+T} = a_i + b_iT$$

where $T$ is the number of time periods from the present time, $t$, to the period for which the forecast was made. The values of $a_i$ and $b_i$ were determined by the following equations:

$$a_i = 2M_{i[1]} - M_{i[0]}$$

$$b_i = \left[\frac{2}{(N-1)}\right] \times \left[ M_{i[1]} - M_{i[0]} \right]$$
2.3 Exponential smoothing method

Exponential smoothing methods are some of the popular methods of forecasting because they are easy to use, require very little computer time and need only a few data points to obtain future predictions. The exponential smoothing method also eliminates the difficulties faced in the regression methods. Exponential smoothing methods have many of the same advantages as moving average technique but require a minimum amount of data storage. The basic exponential smoothing model is

\[
S_t^{[1]} = \alpha X_t + (1-\alpha)S_{t-1}^{[1]}
\]  

or

\[
\text{New estimate} = a(\text{new data}) + (1 - a) \left[ \text{Previous estimate} \right]
\]

\(\alpha\), called the smoothing constant, normally ranges between 0.01 and 0.3. The double exponential smoothing method and triple exponential smoothing method were used in the analysis.

2.3.1 Double exponential smoothing method

Single exponential smoothing cannot deal with non-stationary data. Linear exponential smoothing is an attempt to deal with linear non-stationary data. Its only difference from single exponential smoothing is that it introduces extra formulas that can estimate the trend and subsequently use it for forecasting. To develop an equation that takes account of a linear trend in data, double exponentially smoothed statistics \(S_t^{[2]}\) was calculated. The value of \(S_t^{[2]}\) was determined by the following relation:

\[
S_t^{[2]} = \alpha S_t^{[1]} + (1-\alpha)S_{t-1}^{[2]}
\]  

The initial values of \(S_0^{[1]}\) and \(S_0^{[2]}\) were assumed in these calculations. The initial estimates were not important since relatively large amount of data were available for forecasting. The forecast equation is of the form

\[
y_{t+T} = a_t + b_t T
\]

where \(T\) is the number of time periods from the present time, \(t\), to the period for which the forecast was made. The values of constants \(a_t\) and \(b_t\) were determined by the following relation:

\[
a_t = 2S_t^{[1]} - S_t^{[2]}
\]

\[
b_t = \left[ a_t/(1-\alpha) \right] \times \left[ S_t^{[1]} - S_t^{[2]} \right]
\]

2.3.2 Triple exponential smoothing method

The triple exponential smoothing is also called quadratic exponential smoothing. The triple exponentially smoothed statistics were calculated as follows:

\[
S_t^{[3]} = \alpha S_t^{[2]} + (1-\alpha)S_{t-1}^{[3]}
\]
The initial values of $S_0^{[3]}$ and $S_0^{[4]}$ were also assumed in these calculations. The forecasting equation is of the form

$$y_{t+T} = a_0 + b_1 T + c_1 T^2$$

(15)

where $T$ is the number of time periods from the present time, $t$, to the period for which the forecast was made. The values of constants $a_0$, $b_1$ and $c_1$ are determined by the following relation:

$$a_0 = 3S_0^{[3]} - 3S_0^{[2]} + S_0^{[3]}$$

(16)

$$b_1 = \left[ \frac{\alpha}{2}(1-\alpha)^2 \right] \left[ (6 - 5\alpha)S_0^{[0]} - 2(5 - 4\alpha)S_0^{[2]} + (4 - 3\alpha)S_0^{[3]} \right]$$

(17)

$$c_1 = \left[ \frac{\alpha^2}{2}(1-\alpha)^2 \right] \left[ S_0^{[0]} - 2S_0^{[2]} + S_0^{[3]} \right]$$

(18)

2.4 Autoregressive integrated moving average model

A common approach to forecasting is the Box-Jenkins time series approach, which is used here for forecasting oil demand in India. The objective here was to build an ARIMA model, which adequately represents the data generating processes. The Box-Jenkins method involves the following four-step iterative cycle: model identification, model estimation, diagnostic checking and forecasting with the final model.

In an ARIMA model, the future value of a variable is assumed to be a linear function of several past observations and random errors. That is, the underlying process that generates the time series has the form

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q}$$

(19)

where $y_t$ and $\epsilon_t$ are the actual value and random error at time period $t$, respectively, $\varphi_i$ ($i = 1, 2, \ldots, p$) and $\theta_j$ ($j = 0, 1, 2, \ldots, q$) are model parameters. $p$ and $q$ are integers and often referred to as orders of the model. Random errors, $\epsilon_t$, are assumed to be independently and identically distributed with a mean of zero and a constant variance of $\sigma^2$. The Statistical Package for Social Sciences (SPSS) software was used in this research for oil demand forecast by the ARIMA model.

2.5 Artificial Neural Network model (univariate and multivariate)

In general, ANNs are computational paradigms that implement simplified models of their biological counterpart, biological neural structures. Accordingly, ANNs are characterised by local processing in artificial neuron, i.e. parallel processing, which is implemented by the rich connection pattern between processing elements. The basic building block of the ANN is artificial neuron. The neurons are grouped together in parallel to form layers. The layers are interconnected through the weighting factors. Signals can flow from the input layer through to the output layer in two ways that are unidirectional or bidirectional. In unidirectional connections, the neurons are connected from one layer to the next but not within the same layer. The first and last layers of FFNN are called the input and output layers, and those in between are termed as hidden layers.
The data set involves inputs and outputs of the network. Inputs were past oil consumption data, GNP and population in the case of multivariate ANN model and only past oil consumption data in the case of univariate ANN model and oil demand in various sectors as output in both the cases. Of the total data available, 80% were used for training and the remaining for validation purpose. Once the network was trained, it was used for forecasting future oil demands.

2.5.1 General processing element

The general processing element of the ANN is shown in Figure 2. The individual computational elements that make up the most artifical neural system models are rarely called artificial neurons. They are more often referred to as nodes, units or Processing Elements (PEs). It is not always appropriate to think of the processing elements in a neural network as one-to-one relationship with actual biological neurons.

Figure 2  General processing element of ANN

Each processing element is numbered \( i \). For example, like a real neuron, each processing element has many inputs but only a single output, which can fan out to many other processing elements in the network. The \( i \)th input receives from the \( j \)th processing element, indicated as \( X_j \). Each connection to the \( i \)th processing element has an association with a quantity called weight or connection strength (hereafter, both words are used interchangeably). All these quantities are analogous in the standard neuron model. The output of the processing element corresponds to the strength of the synaptic connection between neurons. In the present models, these quantities were represented as real numbers.

Each processing element determines a net input value based on all its input connections. In the absence of special connections, the net input has been calculated by summing the input values gated (multiplied) by their corresponding weights. In other words, the net input to the \( i \)th unit can be written as

\[
Net_i = \sum_{j} X_j w_{ij}
\]  

(20)

where the index \( j \) runs over all connections to the processing element.
2.5.2 Back Propagation Neural (BPN) network algorithm

The network learnt a predefined set of input–output variable pairs by using two phases of propagate-adopt cycle. After an input has been applied as stimulus to the first layer of neural network units, it is propagated through each upper layer until an output is generated. This output pattern is then compared to the desired output and an error signal is computed for each output unit. The error signals are then transmitted backwards from the output layer to each node in the intermediate layers to organise themselves in such a way that different nodes learn to recognise different features of the total input space. After training, when presented with an arbitrary input pattern that resembles the feature, the individual units learn to inhibit their outputs if the input pattern does not contain the feature that they are trained to recognise.

As the signals propagate through the different layers in the network, the activity present at each upper layer can be thought of as a pattern with features that can be recognised by units in the subsequent layer. The output pattern generated can be thought of as a feature map that provides an indication of the presence or absence of many different feature combinations at the input. The total effect of this behaviour is that the BPN provides an effective means to allowing a computer system to examine data that may be incomplete or noisy and to recognise delicate patterns from the partial input.

The above theory can be summarised and given as a step-by-step algorithmic procedure, which is to be used for developing a C++ code, and developed as a general-purpose user interactive program for the analysis in the present work. The various steps involved in the Feedforward Back Propagation Network (FFBPN) algorithm, which is used for the oil demand forecast, are as follows:

- Apply the input vector to the input units:
  \[ X_p = \left( X_{p_1}, X_{p_2}, X_{p_3} \ldots X_{p_n} \right) \]  
  (21)

- Calculate the net input values to the hidden layer units:
  \[ net_{p_j}^h = \sum W_{ji}^h x_{p_i} \]  
  (22)

- Calculate the outputs from the hidden layers:
  \[ i_{p_j}^h = f_j^h \left( net_{p_j}^h \right) \]  
  (23)

- Move to the output layer. Calculate the net input values to each unit:
  \[ net_{p_k}^o = \sum W_{ki}^o i_{p_j}^h \]  
  (24)

- Calculate the outputs from the output layer:
  \[ O_{p_k} = f_k^o \left( net_{p_k}^o \right) \]  
  (25)

- Calculate the error terms for the output units:
  \[ \delta_{p_k}^o = \left( Y_{p_k} - O_{p_k} \right) f_k^{\prime o} \left( net_{p_k}^o \right) \]  
  (26)
Calculate the error terms for the hidden units:

$$\delta^h_p = f'_o \left( net^h_p \right) \sum \delta^o_j W_{pj}$$  \hspace{1cm} (27)

Update the weights on the output layer:

$$W_{o}^o (t+1) = W_{o}^o (t) + \eta \delta^o_i$$  \hspace{1cm} (28)

Update the weights on the hidden layer:

$$W_{h}^h (t+1) = W_{h}^h (t) + \eta \delta^h_i X_p$$  \hspace{1cm} (29)

where $h$ is the number of the hidden layer in the network;
$j$ is the number of nodes in the hidden layer;
$k$ is the number of nodes in the output layer;
$i$ is the number of nodes in the input layer;
$W$ is the connection strength or weight;
$\delta$ is the error between the actual and the predicted values;
$\eta$ is the learning rate of the network;
$O$ is the output demand calculated by the network;
$X_p$ is the input variable to the neural network;
$Y$ is the actual demand.

In the present oil demand forecast, the network consists of three input neurons, six hidden neurons and one output neuron. So, 18 weights are present between inputs to the hidden layer and six between the hidden and the output layer. Input to the neural network should always be in the range 0.0–1.0. So, all the actual input values should undergo a process of transformation called normalisation. Normalisation is a process similar to interpolation defined as the process of conversion of all inputs into the defined zone. In this present work, intentionally, the range was set between 0.1 and 0.9. This is done to accommodate any further forecast beyond the year 2030. Here 0.1 was made equivalent to the year 1970 data and 0.9 to the year 2030 data. Due to the unavailability of the 2030 demand, some arbitrary value was chosen from regression analysis, whose validation error for the network was calculated. Depending on the validation error and training set results, the 2030 demand was changed (similar to a trial and error procedure). The same process was continued till the validation error was minimised. This kind of input data optimisation is required for the cases where the maximum value in input data is not available. Figure 3 gives the pictorial representation of the multivariate ANN model.

During the training phase, weights were changed continuously until the optimisation of the weights was reached. Modification of the weights depends on the error of the network. This part will be taken care of by the back propagation algorithm. Modification of the weights depends on the learning rate, which is defined as a parameter taking care of the magnitude of change in the values of connection strengths. Weights were modified until the least Mean Square Error (MSE) is reached.
3 Oil demand forecasting

Past oil consumption, GNP and population growth were used for future oil demand prediction. The above-mentioned forecasting methods were used for the prediction of oil demand in various sectors in India. Also, all the forecasting techniques were validated using Mean Percentage Error (MPE) and whichever technique gives less error was used for forecasting. The data used in the study represent the year-wise consumption of oil in different sectors, namely agricultural sector, commercial sector, domestic sector, industrial sector, power sector, transportation sector and other sectors including the total consumption in India. The different forecasting models, namely linear model, exponential model, power model, quadratic model, double moving average method, double exponential smoothing method, triple exponential smoothing method, ARIMA model and ANN model (univariate and multivariate), are listed in Table 1.
Table 1  Comparison of forecasting models for the sector-wise and total oil consumption in India

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Forecasting model</th>
<th>Mean Percentage Error (MPE)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural</td>
<td>( y = 119.25x - 236414.94 ) ( y = e^{76.95 + 0.0091x} ) ( y = 220.09x^{0.8454} ) ( y = 21.09x^2 - 344.49x + 1363.99 )</td>
<td>30 44.7 11.1 21.6</td>
</tr>
<tr>
<td></td>
<td>( y_{x,T} = 6139.25 + 3444.5T ) ( y_{x,T} = 4485.39 + 668.6T ) ( y_{x,T} = 5523.49 + 1654.02T + 93.43T^2 )</td>
<td>17 11 13</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>ARIMA model (1-0-1)</td>
<td>26</td>
</tr>
<tr>
<td>Commercial</td>
<td>( y = 50.42x - 99890.68 ) ( y = 8e^{-124}e^{0.18x} ) ( y = 29.31x^{1.105} ) ( y = 0.86x^2 + 31.54x - 28.42 )</td>
<td>8.5 11.9 5.4 1.32</td>
</tr>
<tr>
<td></td>
<td>( y_{x,T} = 1039.5 + 102T ) ( y_{x,T} = 995.37 + 56.91T ) ( y_{x,T} = 1008.51 + 69.42T + 1.18T^2 )</td>
<td>12 30.8 40</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>12.4</td>
</tr>
<tr>
<td></td>
<td>ARIMA model (1-0-1)</td>
<td>21.2</td>
</tr>
<tr>
<td>Domestic</td>
<td>( y = 230.09x - 455678.78 ) ( y = 6e^{-107}e^{0.127x} ) ( y = 212.62x^{0.5069} ) ( y = 9.25x^2 + 27.13x + 444.42 )</td>
<td>8 13 1.2 2.4</td>
</tr>
<tr>
<td></td>
<td>( y_{x,T} = 5524 + 642T ) ( y_{x,T} = 5197.82 + 371.34T ) ( y_{x,T} = 5395.53 + 559.23T + 17.79T^2 )</td>
<td>12 35.5 46</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>ARIMA model (2-0-2)</td>
<td>2.04</td>
</tr>
</tbody>
</table>
Table 1  Comparison of forecasting models for the sector-wise and total oil consumption in India (continued)

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Forecasting model</th>
<th>Mean Percentage Error (MPE)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial Sector</td>
<td>( y = 470.66x - 925114.44 )</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>( y = 7143e^{0.031} )</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>( y = 6096.9x^{0.279} )</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>( y = 20.91x^2 + 11.88x + 8077.97 )</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 17172 + 220T )</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 17339.85 + 665.96T )</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 17444.69 + 766.04T + 9.44T^2 )</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>ARIMA model (2-0-2)</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>( y = 33.75x - 64432.77 )</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td>( y = 2400e^{0.001x} )</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>( y = 2294.8x^{0.012} )</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>( y = 3.35x^2 - 39.79x + 2632 )</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 4289 + 1263T )</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 3707.81 + 191.94T )</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 4040.78 + 507.96T + 29.97T^2 )</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>ARIMA model</td>
<td>4.1</td>
</tr>
<tr>
<td>Power sector</td>
<td>( y = 1199.9x - 2367570.75 )</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>( y = 2E^{-5}e^{0.064} )</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>( y = 6832.2x^{0.466} )</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>( y = 6.46x^2 + 1345.08x + 6455.32 )</td>
<td>7.9</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 19627.75 + 988T )</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 29052.39 + 367.43T )</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td>( y_{t+1} = 27226.47 - 1364.29T - 164.33T^2 )</td>
<td>25.6</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>ARIMA model (3-0-3)</td>
<td>8.6</td>
</tr>
</tbody>
</table>
Table 1  Comparison of forecasting models for the sector-wise and total oil consumption in India (continued)

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Forecasting model</th>
<th>Mean Percentage Error (MPE)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other sectors</td>
<td>( y = 184.63x - 365182.75 )</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>( y = 1E^{-57} x^{0.0098} )</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>( y = 823.69x^{0.3874} )</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>( y = 24.86x^2 - 361.74x + 2290.08 )</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>( y_{r,T} = 6811.25 + 1222.5T )</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>( y_{r,T} = 5964.06 + 575.66T )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>( y_{r,T} = 6519.92 + 1103.49T + 50.03T^2 )</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>ARIMA model (1-0-1)</td>
<td>5.5</td>
</tr>
<tr>
<td>Total consumption</td>
<td>( y = 2782.57x - 5473332 )</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>( y = 17850e^{0.105} )</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td>( y = 10898x^{0.558} )</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>( y = 87.52x^2 - 105.59x + 21922.89 )</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>( y_{r,T} = 107200 + 1600T )</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>( y_{r,T} = 108292.7 + 4732.43T )</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td>( y_{r,T} = 108867 + 5281.69T + 51.69T^2 )</td>
<td>25.7</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (univariate)</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Artificial neural network model (multivariate)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ARIMA model (1-0-1)</td>
<td>2.1</td>
</tr>
</tbody>
</table>

Figure 4 shows the forecast of oil consumption for the years 2020 and 2030 in India in different sectors such as agricultural sector, commercial sector, domestic sector, industrial sector, power sector, transportation sector and other sectors including the total oil demand. It was found that the consumption of oil is increasing invariably in all the sectors every year. The forecast of oil consumption in India for the years 2020 and 2030 was 3242 and 5435 thousand tonnes, respectively, in the agricultural sector; 2591 and 3650 thousand tonnes, respectively, in the commercial sector; 48,292 and 102,321 thousand tonnes, respectively, in the domestic sector; 42,005 and 57,719 thousand tonnes, respectively, in the industrial sector; 29,549 and 42,179 thousand tonnes, respectively, in the power sector; 38,652 and 42,797 thousand tonnes, respectively, in the transportation sector; and 104,756 and 214,687 thousand tonnes, respectively, in the other sectors. Also, the forecast of total oil consumption in India for years 2020 and 2030 was 415,373 and 720,688 thousand tonnes, respectively.
4 Model validation

Model validation has to be carried out in order to evaluate its accuracy. A comparison of the ANN model (univariate and multivariate) with the time series model and ARIMA model was made for the validation purposes. The predicted oil demand is compared with the actual past oil consumption data and the percentage error was arrived. The results show that forecasting of the oil demand predicted by the ANN model was closer to the real data than that predicted by the other models in most of the cases. The objective was to minimise the MPE and the Root Mean Squared Error (RMSE). The correlation coefficient ($R^2$) is also calculated to determine the optimum forecasting model.

Figure 5  Validation of forecasting models for oil consumption in the agricultural sector
Figures 5–12 represent the validation of both sector-wise consumption and total consumption of oil in India. From Table 1, it is clear that the ANN model (univariate) gives the least MPE in most of the sectors, namely agricultural sector, commercial sector,
domestic sector, industrial sector, other sectors and total oil consumption. The MPE for the ANN model (univariate) for the above-mentioned sectors were 10.8%, 0.64%, 0.45%, 0.8% 1.41% and 0.9%, with the correlation coefficient ($R^2$) of 0.83, 0.75, 0.91, 0.77, 0.62 and 0.88. In the case of the power sector and the transportation sector, the respective double moving average model (MPE: 2.3%) and power model (MPE: 1.1%) were the best models as they hold less MPE.

Figure 8  Validation of forecasting models for the consumption of oil in the industrial sector

Figure 9  Validation of forecasting models for oil consumption in the power sector
Figure 10 Validation of forecasting models for oil consumption in the transportation sector

Figure 11 Validation of forecasting models for oil consumption in the others sector
After validating the different forecasting models, it was concluded that the ANN-based forecasting model can be used for forecasting oil demand in India since it gives the least MPE and RMSE in most of the cases. In the present paper, the forecasting model which gives the least MPE and RMSE was used for forecasting oil demand in India for the years 2020 and 2030.

5 Conclusion

Various forecasting models were developed for the prediction of oil demand in India. After the formulation of forecasting models, model validation was performed to select the best forecasting model. It is found that the ANN-based forecasting model gives the least MPE in most of the cases. Then based on the optimum forecasting model, oil demand in various sectors including the total oil demand was predicted for the year 2020 and 2030 for India. It was found that oil consumption is increasing invariably in all the sectors every year. Also, the forecast of the total oil consumption in India for the years 2020 and 2030 is estimated to be 415,373 and 720,688 thousand tonnes, respectively. This study would be highly useful for the policy-makers for future energy planning in India.

References


Road transport energy consumption in the G7 and BRICS: 1973–2010

Yi-Xuan Gao and Hua Liao*

School of Management and Economics,
Beijing Institute of Technology,
Beijing 100081, China
and
Centre for Energy and Environmental Policy Research,
Beijing Institute of Technology,
Beijing 100081, China
Email: gaoyx1991@gmail.com
Email: hliao@bit.edu.cn
*Corresponding author

Paul J. Burke

Arndt-Corden Department of Economics,
Australian National University,
Canberra, ACT 2601, Australia
Email: paul.j.burke@anu.edu.au

Yi-Ming Wei

School of Management and Economics,
Beijing Institute of Technology,
Beijing 100081, China
and
Centre for Energy and Environmental Policy Research,
Beijing Institute of Technology,
Beijing 100081, China
Email: ymwei@263.net

Abstract: Road transport will account for a large share of developing countries’ future energy demand. This paper reviews the trends in road transport energy consumption in 12 countries (Group of Seven (G7) and BRICS) over the period 1973–2010. We report several stylised facts: road transport energy use and its share in total energy use have been rising; there were large differences in road transport energy use per capita across countries, resulting from differences in country size, resource endowments, fuel prices and other factors; oil accounts for around 95% of road transport energy in the selected countries except Brazil; oil will likely be the dominant road transport energy source in most countries for some years to come but not in the long run; and the use of alternative road transport energy sources is increasing.

Keywords: road transport; energy consumption; historical; G7; BRICS.

Biographical notes: Yi-Xuan Gao is a Graduate Student in the School of Management and Economics, Beijing Institute of Technology, China.

Hua Liao is a Professor in the School of Management and Economics, Beijing Institute of Technology, China.

Paul J. Burke is a Fellow in the Arndt-Corden Department of Economics, Australian National University, Australia.

Yi-Ming Wei is a Professor in the School of Management and Economics, Beijing Institute of Technology, China.

1 Introduction

Transport is a vital sector for any economy and nation. Road transport – the road-based movement of people and freight, whether private or public – is one of the most widely used and popular means of transport and is also a key energy consumer. As incomes rise and urbanisation continues, demand for road transport is increasing. Given this, it is important to document the characteristics, patterns and historical trends of long-term energy consumption for road transport and to compare the experiences of countries at different development levels.

This paper is the first to examine the characteristics of road transport energy demand in the Group of Seven (G7: the United States (USA), Canada, the UK, France, Germany, Italy and Japan) and the five BRICS countries (Brazil, Russia, India, China and South Africa) by descriptive analysis. These countries are important; in 2010, their Gross Domestic Product (GDP; Purchasing Power Parity (PPP), 2005 constant prices), energy consumption and road transport energy consumption accounted for 63.9%, 61.8% and 62.1% of the world’s totals, respectively.

The data we use are from the International Energy Agency (IEA, 2012), International Road Federation (IRF, 2012), World Bank (2014) and US Energy Information Administration (US EIA) (2012). Energy consumption is calculated using the thermal equivalence method. We focus on the almost four-decade period 1973–2010. Russian data are available from 1990. The data for Germany cover both East and West Germany.

Our focus is on the energy used to provide road transportation services, which is a ‘derived demand’ (springing from the demand for transport itself). Our data consider only the energy consumed during transportation rather than energy consumed in the construction of transport infrastructure and equipment.

2 Energy use by the transport sector

In 2010, the global transport sector consumed 2.4 billion tonnes of energy in oil equivalent terms, which equalled 27.2% of global final energy consumption (up 4.1 percentage points since 1973) (IEA, 2012). This makes transport the second-largest energy-consuming sector behind industry.
As seen in Figure 1, the transport share of final energy consumption increased over our period of analysis in all countries in our sample except South Africa. The transport sector accounts for a higher proportion of total final energy consumption in G7 countries than in BRICS countries. IEA statistics on national transport energy use include consumption by road, railway, domestic water-borne, domestic air and pipeline-based transport activities. International aviation and marine transport are not included in national-level data.

Figure 1  Structural variation in total final energy consumption

Notes: World Bank country abbreviations used throughout: BRA = Brazil; CAN = Canada; CHN = China; DEU = Germany; FRA = France; GBR = United Kingdom; IND = India; ITA = Italy; JPN = Japan; RUS = Russia; USA = United States; ZAF = South Africa.

In the G7 countries, industrial energy consumption has levelled off as these countries have moved to more service-based economies and have achieved energy efficiency improvements. At the same time, increasing demand for transport (and more comfortable transport, for example by private vehicles) has seen the amount of energy used in transport increase, despite energy efficiency improvements in transport also. As a result of these two phenomena, the transport share of total final energy use has risen.

In 2010, the USA’s transport energy consumption reached 580 million tonnes of oil equivalent, accounting for 38.9% of final energy consumption (an increase of 7.4 percentage points from 1973). Canada’s transport energy consumption was 60 million tonnes of oil equivalent, equalling 30.4% of final energy consumption (an increase of 4.8 percentage points since 1973). Transport energy consumption has reached approximately 40 million tonnes of oil equivalent in each of France, Italy and the UK, nearly 30% of their final energy consumption (increases of approximately 10 percentage points since 1973). Germany and Japan’s year 2010 transport energy consumption was 50 million and 80 million tonnes of oil equivalent, accounting for 23.7% and 23.4% of final energy consumption (increases of 8.4 and 6.3 percentage points since 1973). Transport, rather than industry, has become the largest energy-consuming sector in France, Italy, the UK and the USA.
In the BRICS countries, industrial energy use has continued to increase, but transport energy consumption has increased even more rapidly. In China and India, transport remains a relatively small, although growing, contributor to total energy consumption. By 2010 the transport share of final energy consumption reached 11.5% in China (up from 4.5% in 1973) and 12.1% in India (up from 9.2%). Transport guzzled 170 and 60 million tonnes of oil equivalent in the two countries in 2010, respectively.

Transport-sector energy consumption in Brazil reached 70 million tonnes of oil equivalent in 2010. The transport sector’s share of total final energy consumption in Brazil – at 26% in 1973 and 33% in 2010 – was higher than in the other BRICS countries. This is as a result of factors including Brazil’s large land area, relatively low population density and relatively low residential energy consumption (only 11% of total energy consumption in 2010).

South Africa is the only country in our sample to have had a declining share of transport in final energy over the period. The proportion of final energy use attributed to the transport sector in South Africa decreased over the period 1973–1990, followed by a slight increase from 1990 to 2010. The reason for this mainly lies in South Africa’s complicated history. Due to government policies in the apartheid era, its per capita real GDP showed an overall downward trend over the period from 1972 until the abolition of racial segregation in 1994.

3 Road transport

3.1 Road transport energy consumption versus total transport energy consumption

Road transport usually makes up the majority of total transport energy consumption, a share that is typically increasing (Figure 2). From 1973 to 2010, global road transport energy consumption rose from 0.7 billion to 1.8 billion tonnes of oil equivalent, an average annual growth rate of 2.6% per annum. Over the same period, global final energy consumption and GDP grew at average annual rates of 1.7% and 3.2%, respectively.
In G7 countries, the share of road transport in domestic transport energy use exceeded 80% by 2010, having gradually increased over the last 40 years. Germany is the most road transport energy-dependent (94.7% of transport energy), followed by France at 93.8%. In both Italy and the UK, the proportions were 92.7%.

From 1973 to 2010, the road share of transport energy use in China, India and South Africa increased substantially. China and India both doubled their shares (China: from 39.6% to 77.3%; India: from 42% to 88%). South Africa’s road share of transport energy use increased from 66.7% to 90.8% over the period. The proportion of rail transport energy use in total transport energy use in these three countries decreased substantially during this period (China: from 42.3% to 6.9%; India: from 55.0% to 6.7%; South Africa: from 31.4% to 2.6%) as road transport has become increasingly dominant. Russia had the lowest proportion of road transport energy consumption (49.6%) in 2010, up from 44.9% in 1990. The reasons for this low share lie in Russia’s cold climate (which dissuades some road-based travel) and its widespread use of pipeline-based transport of oil and natural gas (with a transport energy consumption proportion of 36.6%).

3.2 National stories on road-transport energy use

Figure 3 shows that, since 1973, national road transport energy consumption of each country in our sample has increased, but at different rates.

![Figure 3 Total road transport energy consumption](image)

Notes: toe is ‘tonnes of oil equivalent’.

Owing to its large land area, large population, high income level, abundant resources and profligate consumption habits, the USA consumes far more road transport energy than other countries. In 2010, the total road transport energy consumption of the USA was 520 million tonnes of oil equivalent, about 13.7 times that of the UK, 7.5 times that of Japan and 3.4 times that of China. The USA accounts for 28.8% of global road transport energy consumption; all the other countries in our sample combined account for 34.1%.

France, Germany, Italy, the UK and Japan display generally similar road transport energy consumption trends due to their similarities in states of development and geographical conditions. Since 1973, the road transport energy consumption of each of these countries has increased. Over the period 1973 to 1990, their road transport energy
average annual growth rates were about 3%. In the period since 1990, these average annual growth rates fell to less than 1%. Japan’s total road transport energy consumption (69 million tonnes of oil equivalent) was higher due to its larger population.

China’s rapid economic development has seen its annual road transport energy consumption grow from 6.5 million tonnes of oil equivalent in 1973 to 150 million tonnes of oil equivalent in 2010 (an annual average growth rate of 8.9%). Other BRICS countries such as India and Brazil have also shown rapid growth in road transport energy use. From 1973 to 2010, road transport energy consumption in India increased from 5.7 million tonnes of oil equivalent to 48.8 million tonnes of oil equivalent, an average growth rate of 6.0% per annum. South Africa’s use of energy in road transport increased from 6.4 million tonnes of oil equivalent to 12.0 million tonnes of oil equivalent, an average annual growth rate of 3.7%.

In terms of road transport energy consumption per unit land area, Japan ranks first, using 211 tonnes of oil equivalent/km² in 2010 (IEA, 2012; World Bank, 2014; and authors’ calculation). The UK, Germany and Italy also have relatively high values (171, 152 and 131 tonnes of oil equivalent/km², respectively). The (less densely populated) USA and Canada consume only 63.8 and 6.5 tonnes of oil equivalent/km², respectively. For BRICS countries, the values are far less than for G7 countries (other than Canada). The per unit land area road transport energy consumption for China and India are both 18.7 tonnes of oil equivalent/km²; South Africa, Brazil and Russia consume 10.9, 8.3 and 5.9 tonnes of oil equivalent/km², respectively.

3.3 Road transport is gradually becoming the main oil consumer in all countries

The road transport share of total final oil consumption has increased, both globally and in our countries of focus (Figure 4). This share is particularly high in G7 countries, where more oil is consumed in road transport than in all other final-consuming sectors combined (except Japan). Among BRICS countries, this share is also high in South Africa, Brazil and Russia. China and India currently use a relatively large share of their (final-use) oil in their industrial sectors, but are most likely on track to a similar situation of road-transport sector dominance in national final oil demand. In the long run this situation may change, but historical trends are clearly towards a growing share of oil being used in road transport (Keshavarzian et al., 2012; Wang et al., 2014). There has been falling dependence on oil in some other sectors, such as electricity generation (Burke, 2010).

3.4 Per capita road transport energy consumption versus per capita GDP: a longitudinal perspective

Over time, per capita road transport energy consumption has generally been positively correlated with per capita GDP (Figure 5). In G7 economies, however, there has been some lessening in the extent to which per capita road transport energy consumption increases with per capita GDP as the road transport sector has matured. The GDP per capita at which per capita road transport energy consumption has reached a local peak has differed across countries: this level was around $25,000 and $23,000 in the USA and Canada (2005 constant prices, PPP and similarly hereinafter), while in the EU countries and Japan it was about $30,000. Per capita road transport energy consumption has in some instances (e.g., USA, Canada) since begun to increase again, although the USA
again experienced declines during its economic recession in the final years of the sample period. It is certainly too early to conclude that these countries are near ‘peak energy use in the road transport sector’. Figure 4 also shows that road transport energy use can show short-run declines, especially during recessions and/or times of high oil prices.

Figure 4  The proportion of road transport oil consumption in final oil consumption in 1973, 1990 and 2010 (see online version for colours)

Notes: Oil is also used in some energy transformation processes, such as the generation of electricity. This oil is not included in ‘final consumption’; it is instead part of the broader concept named ‘primary consumption’. Data are not available for Russia in 1973.

Figure 5  Per capita road transport energy consumption vs. per capita GDP, 1971–2010 (see online version for colours)

Notes: GDP is from IEA (2012). kgoe is ‘kilogrammes of oil equivalent’.
From 1973 to 2010, the per capita road transport energy consumption of China and India increased at average annual rates of 8% and 4%, respectively. Due to the social and economic impacts of collapse of the Soviet Union in 1991, Russia firstly saw a decrease and then rapid growth in per capita GDP and per capita road transport energy consumption during the period from 1990 to 2010. The per capita GDP and per capita road transport energy consumption of South Africa did not show an increasing trend until 1994 (attributable to historical reasons related to the end of apartheid). It can be seen from Figure 5 that very large increases in BRICS’ road transport energy use will occur if they are to follow the path of G7 countries.

Table 1 lists the elasticity of per capita road transport energy consumption to per capita GDP for each country for three time periods from 1960 to 2010. 1973 is used as a break because of the Arab-Israeli war and oil price shock of 1973. 1990 is used as a break for two reasons: (a) the data for Russia start from 1990; (b) the Gulf War broke out in 1990 and the Soviet Union collapsed in 1991, which both had implications for the global economy and energy markets. The ‘road transport energy: income’ elasticity of the G7 countries exceeded 1.0 in the period before 1973 and in some instances exceeded 2.0. Rapid economic development of these countries was associated with even more rapid increases in road transport energy use.

<table>
<thead>
<tr>
<th></th>
<th>USA</th>
<th>CAN</th>
<th>GBR</th>
<th>FRA</th>
<th>DEU</th>
<th>ITA</th>
<th>JPN</th>
<th>BRA</th>
<th>RUS</th>
<th>ZAF</th>
<th>CHN</th>
<th>IND</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960–1973</td>
<td>1.2</td>
<td>1.2</td>
<td>2.2</td>
<td>1.7</td>
<td>2.5</td>
<td>2.0</td>
<td>1.6</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1973–1990</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.0</td>
<td>0.8</td>
<td>–</td>
<td>−0.4</td>
<td>0.8</td>
<td>2.3</td>
</tr>
<tr>
<td>1990–2010</td>
<td>0.2</td>
<td>0.7</td>
<td>−0.1</td>
<td>0.1</td>
<td>−0.1</td>
<td>0.6</td>
<td>0.3</td>
<td>1.5</td>
<td>−0.3</td>
<td>0.6</td>
<td>1.0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Notes: We used the straightforward approach of using the ratio of the two average growth rates for the specific periods. Because we have not controlled for prices and other variables, these elasticities are for illustrative purposes and represent general associations.

From 1973 to 1990, this elasticity reduced in all G7 countries, suggesting a degree of decoupling of road transport energy growth from growth in economic activity. The road transport energy: income elasticities of the USA and Canada were close to zero during 1973–1990, while those of the other countries were c. 1.0.

After 1990, owing to changes in the means by which economic growth was manifested, this elasticity was lower than 1.0 (even negative) in G7 countries. The road transport energy: income elasticities of China and Brazil were below 1.0 before 1990; after 1990, as structural change drove robust growth of road transport, they were 1.0 or higher.

### 3.5 Explanations of differences between countries

The per capita road transport energy consumptions of the USA and Canada were about 1.5 tonnes of oil equivalent in 2010, vastly higher than those of the other countries. Per capita road transport energy consumption in the EU countries and Japan was about 0.6 tonnes of oil equivalent. Brazil, China and India had per capita road transport energy consumptions of 0.3, 0.1 and 0.04 tonnes of oil equivalent, respectively. The sizeable differences between countries are largely driven by the stage of development; countries
with high per capita GDP tend to have higher per capita road transport energy consumption. Per capita road transport energy consumption also varies due to factors such as relative resource abundance, vehicle ownership levels, consumption habits, pricing and other policies and infrastructure availability (Liu et al., 2006; Burke and Nishitateno, 2013).

As might be expected, ownership of four-wheeled motor vehicles is closely associated with per capita road transport energy consumption. Figure 6 shows the relationship between ownership of these vehicles per thousand people and per capita road transport energy consumption for an expanded sample of 84 countries in 2008. The size of the bubbles reflects total vehicle ownership in each country. Naturally enough, higher vehicle ownership levels are associated with higher per capita road transport energy consumption. In 2010, four-wheeled motor vehicle ownership per thousand people exceeded 600 in Canada, France, Italy and Japan: in the USA, it exceeded 800. In contrast, four-wheeled motor vehicle ownership per thousand people is much lower in BRICS countries (e.g., China and India: 37 and 17, respectively) (IRF, 2012).

Figure 6  Four-wheeled motor vehicle ownership per thousand people vs. per capita road transport energy consumption in 2008

Notes: The area of each bubble is proportional to total four-wheeled motor vehicle ownership. Four-wheeled motor vehicle ownership is from the IRF (2012). 84 countries are included.

The number of high energy-consuming vehicles, such as sports utility vehicles, light vans and light trucks has grown rapidly in the USA (Du et al., 2013). This is a key reason why the per capita road transport energy consumption of the USA is higher than in other countries. A recent documentation of how vehicle ownership evolves as economies develop is provided by Nishitateno and Burke (2014).

Resource endowments also appear to influence per capita road transport energy consumption. The USA and Canada are rich in fossil fuel resources: by the end of 2010, their proven crude oil reserves were 2.5 and 28.2 billion tonnes, respectively (US EIA, 2012). Perhaps because of their relatively abundant energy supplies, among other reasons, the USA and Canada have lower fuel tax rates, meaning their fuel prices are relatively low and their consumption is high (Burke and Nishitateno, 2013). France, Germany and Italy are lacking in fossil fuel resources: by the end of 2010, their proven...
crude oil reserves were 10 million, 40 million and 60 million tonnes, respectively (US EIA, 2012). Fossil fuel reserves are extremely deficient in Japan, with proven crude oil reserves of only 6 million tonnes in 2010 (US EIA, 2012). These resource-limited countries have had higher fuel taxes and implemented generally tighter energy-saving measures (Burke and Nishitateno, 2013). Nevertheless, almost all countries have adopted some policies to promote the development of public transport and/or encourage the use of fuel-efficient and/or clean-energy vehicles in an attempt to improve energy efficiency (Wei et al., 2010).

As shown in Figure 7, European and Japanese fuel prices are much higher than those in the USA and Canada. The price differences arise mostly from the different fuel taxes collected by their governments (as well as international transport costs). The high fuel taxes in European countries and Japan reduce road transport energy consumption to some extent, with estimates of the long-run fuel price elasticity of road transport fuel demand typically falling in the range –0.2 to –0.8 (Sterner, 2007; Dahl, 2012; Havranek et al., 2012; Burke and Nishitateno, 2013). There is also evidence that the effects of gasoline prices flow through to other road-sector outcomes, such as the number of road deaths: countries with higher gasoline prices tend to have fewer road fatalities, holding other factors constant (Burke and Nishitateno, 2014).

Figure 7  Gasoline and diesel prices in 2010


Infrastructure availability is also an important factor influencing road transport energy consumption. As shown in Figure 8, per capita road transport energy consumption is positively related to per capita road distance across an expanded sample of 78 countries in 2010. The road networks of Canada and the USA stretched 29.4 and 21.1 metres per capita, respectively, in 2010, far longer than other G7 or BRICS countries due to their large land areas, high levels of economic development and road-prioritising transport development histories. China’s road distance per capita was only 3.0 metres in 2010 (up from 0.02 metres in 1975). The road distance per capita was 3.7 in India in 2010 (up from 2.0 metres in 1973).
4 Structural changes in road transport energy

4.1 Oil is still dominant, but its share in road transport energy is declining

In 1973, the oil share of road transport energy consumption exceeded 99% in the 12 countries we are studying. By 2010 this share had reduced in most countries, although Japan and South Africa had shown little change. In Brazil, the oil share of road transport energy consumption had fallen to only 75% in 2010. The fall in this share was mostly as a result of policies to encourage the substitution of oil with biomass energy: by 2010, Brazil’s biomass energy share in road transport reached 22%. Biomass energy utilisation also increased in France, Italy and the USA, to 5.9%, 4.1% and 4.7%, respectively. Use of natural gas also increased, with the natural gas utilisation ratios of China and India both rising from zero in 1973 to 4.7% in 2010.

Given the pressures on the environment and on fossil fuel resources in general, the proportion of clean energy in final energy consumption will likely increase in the future (Liao and Wei, 2010). A number of countries have adopted policies that promote the uptake of clean energy sources, such as fuel taxes, subsidies for biomass, fuel economy standards and tax incentives to purchase electric cars. As a result of some of these policies, pressure from rising oil prices and ongoing technical progress, the vehicle fuel efficiency ratings of new vehicles have generally been improving in almost all countries (Cuenot and Körner, 2013).

Sales of Electric Vehicles (EVs) and non-plug-in Hybrid-Electric Vehicles (HEVs) grew strongly in 2011 and 2012, but their market shares are still small (IEA, 2013). For example, in 2012, HEVs only made up 1.5% of global new-vehicle sales (IEA, 2013). This may well change over the next decade or two. Currently, however, gasoline and diesel generally remain cheaper road transport options than the alternatives. While
alternative technologies are improving quickly, ongoing improvements in the fuel economy of gasoline and diesel-powered vehicles make the task of achieving mass sales of alternative technologies (e.g., EVs) more challenging. Most affordable EVs still do not run far without being recharged, although advances are ongoing. There is a need for a roll-out of more charging facilities to see broader uptake. Given the challenges associated with alternative energy sources, oil will likely keep its dominant position in most countries in the medium term. In the long run, however, pollution concerns and resource constraints mean it is likely that alternative technologies such as EVs will come to dominate the market (Huo et al., 2012; He and Chen, 2013). If their costs continue to decrease, in the long run a mass transition to non-oil powered vehicles appears inevitable.

4.2 The rise of diesel

The main oil products used in road transport are gasoline and diesel. As shown in Figure 9, the relative importance of diesel and gasoline varies. The proportion of diesel consumption increased in all of the countries due to the increasing popularity of diesel vehicles given their higher energy efficiency and the lower taxes on diesel fuel in most countries (Harding, 2014). Some countries, particularly in Europe, have also had policies to encourage the use of diesel vehicles (Wallington et al., 2013).

Figure 9  Energy sources used in road transport

In 1973, gasoline consumption in the road sector exceeded diesel consumption in all of the countries in our sample except India. By 2010 the proportion of diesel consumption had increased (also except in India). In the UK, France, Germany and Italy, diesel is now the main road transport fuel. From 1973 to 2010, China’s use of diesel for road transport increased sizeably due to rapid growth in freight transport. Gasoline and diesel now rank as equally important as road transport fuel sources in China.
5 Conclusions

Transport is a key energy consuming sector and road transport is the major part thereof. With continued increases in income levels, in most countries the shares of transport energy in total final energy consumption and of the road transport sector in transport energy consumption will likely continue to increase for some years to come.

The per capita road transport energy consumptions of different countries vary widely, with the USA and Canada far exceeding the other G7 countries. In addition to economic development, per capita road transport energy consumption is affected by factors such as fuel tax policies, the abundance of resources, road infrastructure and residential consumption habits (e.g., vehicle ownership decisions). For a variety of reasons, road transport energy consumption growth is in a decelerating phase in developed countries. It is not possible to conclude that these countries are near ‘peak energy use in the road transport sector’, however.

Oil has been the main source of road transport energy. This situation may persist in the medium term, despite advances in alternative technologies such as electric cars. The proportion of diesel consumption is increasing. The market share of clean, renewable energies in road transport energy remains small, but they are likely to play a larger role in the future (perhaps via clean electricity generation). The extent to which EVs and HEVs will help to reduce emissions depends partly on the uptake of low-carbon generation sources in the electricity sector, the prospects for which are generally improving (Cong, 2013; Cong and Shen, 2014).

As a rapidly developing country, China’s road sector has been expanding quickly due to urbanisation, infrastructural improvement and increasing personal incomes. Oil will likely be China’s main source of road transport energy consumption for some years to come, although pollution and resource considerations suggest that it would perhaps be sensible for China to not follow the development paths of the USA or Canada when it comes to road transport energy use. Fortunately vehicle fuel economy and alternative vehicle technologies are improving and there are numerous policy options available for reducing negative externalities from the road transport sector. These include externality pricing of emissions and congestion (Burke, 2014), public transport initiatives, support for research into and adoption of new-energy vehicles and vehicle energy efficiency initiatives. China has already implemented policies of sizeable ambition, including new-vehicle fuel economy standards, rail investments and programmes for the adoption of EVs (Gong et al., 2013).

Acknowledgements

We are thankful for financial support from the ‘Strategic Priority Research Program’ of the Chinese Academy of Sciences (No. XDA05150600), National Natural Science Foundation of China (No. 71322306, 71273027 and 71020107026) and Program for New Century Excellent Talents in University of Ministry of Education of China (No. NCET-13-0040).
References


Note

1 The data are from the IEA (2012) unless otherwise indicated.
A bibliometric analysis of energy poverty research: results from SCI-E/SSCI databases

Kang Li*, Su-Yan Pan and Yi-Ming Wei

School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China
and
Center for Energy and Environmental Policy Research, Beijing Institute of Technology, Beijing 100081, China
Email: likingno.2@163.com
Email: pansuyan@sina.com
Email: wei@bit.edu.cn
*Corresponding author

Abstract: Based on the online version of SCI-E and SSCI databases, this paper employed bibliometrics method to analyse the scientific publications on energy poverty from 1981 to 2013. By analysing the basic characteristics of energy poverty publications, this paper found energy poverty is an interdisciplinary subject, and identified the most productive journals, authors and institutions, the most cited publications and the collaboration relationship among authors, institutions and countries. By frequency analysis of keywords, five hot research topics of energy poverty were summarised. They were fuel access, energy efficiency, physical health, electrification and income poverty. This paper compared the academic research level of the main countries by assessment model, and found that the academic research level on energy poverty of developed countries was much higher than that of developing countries, and England had the total predominance in this field.

Keywords: energy poverty; bibliometrics; assessment model; basic characteristics; hot topics; frequency analysis.


Biographical notes: Kang Li is currently a PhD candidate at the School of Management and Economics and the Center for Energy and Environmental Policy Research, Beijing Institute of Technology (BIT), China. Her research interests focus on energy poverty, energy equity and so on.

Su-Yan Pan is currently a PhD candidate at the School of Management and Economics and the Center for Energy and Environmental Policy Research, Beijing Institute of Technology (BIT), China.

Yi-Ming Wei is the Dean and Professor of the School of Management and Economics, Beijing Institute of Technology (BIT), China. He is also the Founding Director of the Center for Energy and Environmental Policy
1 Introduction

The world faces many important challenges, and energy poverty may be one of them. Energy poverty could restrict further social and economic development, undermine existing achievements, spoil any attempts at achieving energy equity, aggravate environmental pollution and harm physical health (Birol, 2007; Wei et al. 2012). Energy poverty and fuel poverty have all been addressed, and both of them descript the problems of households’ energy consumption.

Fuel poverty has been the concern of some campaigners since the 1970s or earlier. In the early 1980s, fuel poverty became a serious political issue, and became the subject of formal government legislation in UK. Most of fuel poverty researches took residents of England, Ireland and Scotland as research targets. Lewis (1982) first defined the concept as the inability to afford adequate warmth in the home; Boardman’s (1991) definition that if a householder needed to spend more than 10% of their income on total household fuel costs to achieve a satisfactory indoor temperature regime, then they were classed to be in fuel poverty. This definition was refined and then officially adopted by the UK government in 2001 (DEFRA and DTI, 2001); Boardman (2010) later defined fuel poverty thus: households could not “afford adequate services … clearly demonstrated when the home is cold or fuel debts accumulate”. Hills (2011) proposed a new definition of fuel poverty in a 2013 review called the ‘Low Income High Cost’ (LIHC) measure being those householders who would need to spend more on fuel costs than the median level and having done so the result would mean that their residual income would put them below the official poverty line, and it was also officially adopted by the UK government (DECC, 2013).

According to Bouzarovski (2012), there were over 50 million Europeans who as of 2012 cannot maintain a satisfactory indoor heating regime, and the percentage in New Zealand was about 14% (Lloyd, 2006). The UK government suggested that in 2011 there were about 4.5 million fuel poor in UK (7.1%). The annual total residential energy consumption of Europe, New Zealand and UK is 29.29 GJ per capita, 30.24 GJ per capita and 15.48 GJ per capita, respectively, and the consumption of modern fuels accounts for more than 90%.

Fuel poverty is caused by many reasons, such as energy-inefficient housing and heating system (Healy and Clinch, 2002; Shortt and Rugkasa, 2007) and low incomes (Healy and Clinch, 2004; Saunders et al., 2012; Wright, 2004). Fuel poverty is thought to obstruct the development of health, well-being and social equity. In poor-quality buildings, households must consume more energy to reach the satisfactory indoor temperature standards (Healy and Clinch, 2002). The fuel poor are more vulnerable to high fuel prices, and fuel poverty may constrain household income and expenditure into a vicious circle. Fuel poverty could give rise to exacerbation of social isolation, and may jeopardise educational achievement (Roberts, 2008). In colder climates, living in a comfortably heated home is commonly viewed as protective for human health (Liddell
A bibliometric analysis of energy poverty research

and Morris, 2010), especially for the elderly and young children. Living in the cold, poorly equipped houses for long could increase mortality rates (Wilkinson et al., 2007) and have a detrimental influence upon physical health (Liddell and Morris, 2010).

Energy poverty has often been defined as a lack of access to modern energy services, and most of the energy poverty publications focus on developing countries. UNDP first defined energy poverty as “the absence of sufficient choice in accessing adequate, affordable, reliable, high-quality, safe, and environmentally benign energy services to support economic and human development” (UNDP, 2000). IEA (2002) defined energy poverty as households who cannot get access to electricity and commercial energy, and in 2010 as households who cannot get access to electricity and clean cooking facilities (IEA, 2010). Most of the energy poverty researchers have accepted the concept of IEA (Sagar, 2005; Birol, 2007; Bhide and Rodriguez, 2011; Sesan, 2012; Kaygusuz, 2011), and some expanded the energy poverty concepts. Sovacool et al. (2012) recommended that mobility and mechanical power should be included in energy poverty, and the energy poor have also been defined as households who cannot meet their basic energy needs by estimating a minimum limit of energy consumption (Pereira et al., 2011).

There are about 1.3 billion people in the world who do not have access to electricity in 2011 (IEA, 2013), with most of the electricity-deprived population living in the developing world, mainly in Africa, developing Asia and Latin America. The lowest levels of electrification rate of the world are currently in sub-Saharan Africa, where only 31% of the population have access to electricity. It has been estimated by the IEA that the number of electricity-deprived population in the world will fall by only around 0.2 billion during the next 20 years, without any new policies to address this problem. Even then, it will still account for around 15% of the world’s population (IEA, 2010). Access to electricity is the result of policy-driven electrification projects in most cases. China’s electricity industry experienced significant growth over the past three decades, because of strong government intervention (Chen and He, 2013). Most of the households under energy poverty, who rely on traditional use of biomass for cooking, are living in developing countries, with 32% of them in India, 24% in sub-Saharan Africa and 16% in China. The population using traditional devices of biomass of the world is estimated by the IEA to increase slightly from 2.7 billion in 2010 to 2.8 billion in 2030 (IEA, 2010).

Energy poverty is considered as one of the most important topics related to development. Access to electricity is thought important to generate employment opportunities, to elevate standards of education, to improve health status and to facilitate sustainable development. It has been shown that lack of electricity can exacerbate poverty and contributes to its persistence (Pereira et al., 2010). It has also been shown that household electrification can lead to a significant reduction of energy poverty and a consequent improvement in energy equity (Pereira et al., 2011). In addition it is well known that the use of traditional biomass in inefficient ways has a detrimental influence upon health, restricts the development of education, aggravates local ecological damage and can be the cause of a vicious cycle of a lack of economic and social development (Birol, 2007; Kaygusuz, 2011). There is a strong link between income and access to electricity, as lack of electricity precludes most industrial activities and the jobs they create. Expanding electricity access often directly contributes to the goal of eradication of poverty. Regular access to electric energy is a key issue for the rural economic development and for the alleviation of poverty (Pereira et al., 2011). Lighting could bring about better learning environment for students to study outside of daytime, and lower the risk of deterioration of their eyesight (WHO, 2006). Access to electricity also gives an
opportunity to build an education system based on multimedia. Indoor lighting allows doctors to treat patients outside of daylight hours, advanced medical facilities need access to an adequate electricity supply to operate and, in addition, access to multimedia based on electricity promotes efficient information exchange among doctors. People exposed to indoor air pollution from burning biomass by traditional facilities are more likely to suffer from inflammation of the airways and lungs and impaired immune response. Indoor air pollution from burning solid fuel is one of the top ten global health risks (WHO, 2006). Collecting and managing traditional biomass is widely viewed as women’s responsibility (Kaygusuz, 2011), and cooking on the traditional use of biomass makes women more vulnerable to indoor air pollution. Access to modern energy services is one path to ensure gender equity.

Both energy poverty and fuel poverty focus on energy consumption of the residential sector, and they constitute the existing research system of energy poverty (Li et al., 2014). So this paper analysed publications of energy poverty and fuel poverty, and called them energy poverty uniformly.

2 Materials and method

The databases used for this paper were the SCI-E and SSCI databases from Web of Science, and they are the most frequently used indexed databases in bibliometrics analysis. Searches were done under the search ‘Topic’ section, using four variant keywords: ‘energy poverty’ or ‘energy poor’ or ‘fuel poverty’ or ‘fuel poor’, from all documents published from 1981 to 2013 through a cut-off date of 21 January 2014. The number of original publications found using the search parameters for this period was 318. In order to ensure the publications used for analysis referred to energy poverty which was mentioned above, this paper again selected samples by artificial selection and the final sample included 269 publications.

For each publication, all information relevant to the analysis – author(s), editor(s), title, source, addresses, times cited, keywords, abstract, language and Web of Science category – was exported, and the analysis was performed by using the ‘analyse’ function of the Web of Science-based software and Bibexcel software.

3 Results

3.1 Publication trends

The number of publications showed energy poverty research has stepped into the stage of initial development. Figure 1 shows the number of published articles of the whole world by year of publication. From 1993 to 2006, the energy poverty research was still in its infancy, and from 2007 to 2013 it showed a general increasing tendency in the number of published articles. The average annual growth rate was 48.5%, with some fluctuations.

Figure 1 also shows the distribution of publications of top five countries with the highest number of publications. Among the 269 publications analysed in this paper, 89 (33.0%) were from England, more than other 43 countries, which have published energy poverty publications. USA published 42 energy poverty articles, accounting for 15.6%, and the first article of USA was published in 2005. There were three publications on energy poverty from China, less than those of other developing countries such as India and Brazil.
3.2 Subject categories and journals

Energy poverty is a multidisciplinary subject, including 38 subject categories, and the three most common categories were energy economics (38%), environmental sciences (28%) and engineering (10%) (see Figure 2).
In total, 269 publications referring to energy poverty were published by 101 journals. Table 1 lists the information of the ten journals with the highest number of publications, including the number of publications, the percentage of 269 articles, the IF and the location of the journal. *Energy Policy* has published 81 articles related to energy poverty, accounting for 30.1% of the publications considered. *Renewable & Sustainable Energy Reviews* has published 13 considered articles, accounting for 4.8%. Both of these journals are published in England.

**Table 1** Most productive journals in energy poverty research

<table>
<thead>
<tr>
<th>Journal</th>
<th>NP</th>
<th>NP (%)</th>
<th>IF</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Policy</td>
<td>81</td>
<td>30.1%</td>
<td>2.7</td>
<td>England</td>
</tr>
<tr>
<td>Renewable &amp; Sustainable Energy Reviews</td>
<td>13</td>
<td>4.8%</td>
<td>5.6</td>
<td>England</td>
</tr>
<tr>
<td>Energy for Sustainable Development</td>
<td>10</td>
<td>3.7%</td>
<td>2.2</td>
<td>Netherlands</td>
</tr>
<tr>
<td>Applied Energy</td>
<td>9</td>
<td>3.3%</td>
<td>4.8</td>
<td>England</td>
</tr>
<tr>
<td>Energy and Buildings</td>
<td>6</td>
<td>2.2%</td>
<td>2.7</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Energy</td>
<td>5</td>
<td>1.9%</td>
<td>3.7</td>
<td>England</td>
</tr>
<tr>
<td>Environment and Planning A</td>
<td>5</td>
<td>1.9%</td>
<td>1.9</td>
<td>England</td>
</tr>
<tr>
<td>Housing Studies</td>
<td>5</td>
<td>1.9%</td>
<td>0.7</td>
<td>Scotland</td>
</tr>
<tr>
<td>Oil Gas Journal</td>
<td>5</td>
<td>1.9%</td>
<td>0.2</td>
<td>USA</td>
</tr>
<tr>
<td>Renewable Energy</td>
<td>5</td>
<td>1.9%</td>
<td>3.0</td>
<td>England</td>
</tr>
</tbody>
</table>

Notes: NP: number of publications; NP (%): percentage of publications considered; IF: impact factor.

### 3.3 Authorship

The total number of authors related to the accounted 269 publications was 518, and 83.0% (n = 430) and 11.1% (n = 58) of them were co-authors in only one and two articles, respectively. Table 2 displays the most productive authors of energy poverty research. Singapore and Austria contributed two each. Sovacool, who worked at the National University of Singapore and the Vermont Law School in succession, published 24 of the considered articles.

**Table 2** Most productive authors in energy poverty research

<table>
<thead>
<tr>
<th>Author</th>
<th>NP</th>
<th>Country</th>
<th>C</th>
<th>CPP</th>
<th>H index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sovacool, B.K.</td>
<td>24</td>
<td>USA</td>
<td>141</td>
<td>5.9</td>
<td>8</td>
</tr>
<tr>
<td>D’agostino, A.L.</td>
<td>6</td>
<td>Singapore</td>
<td>56</td>
<td>9.3</td>
<td>4</td>
</tr>
<tr>
<td>Pachauri, S.</td>
<td>6</td>
<td>Austria</td>
<td>93</td>
<td>15.5</td>
<td>4</td>
</tr>
<tr>
<td>Bambawale, M.J.</td>
<td>5</td>
<td>Singapore</td>
<td>35</td>
<td>7.0</td>
<td>3</td>
</tr>
<tr>
<td>Bazilian, M.</td>
<td>5</td>
<td>Austria</td>
<td>28</td>
<td>5.6</td>
<td>3</td>
</tr>
<tr>
<td>Healy, J.D.</td>
<td>5</td>
<td>Ireland</td>
<td>163</td>
<td>32.6</td>
<td>5</td>
</tr>
<tr>
<td>Liddell, C.</td>
<td>5</td>
<td>Northern Ireland</td>
<td>19</td>
<td>3.8</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: NP: number of publications; Country: the country of the first unit of the latest publication; C: number of citations received; CPP: citations per publication; H index: based on these 269 publications.
3.4 Institutions

The 269 articles were from 241 institutions, among which the National University of Singapore (13), the Vermont Law School (11) and the University of Oxford (10) were the top three research institutions ranked by productivity. Table 3 lists the top 15 most productive institutions, with England contributing seven of them and none from developing countries.

<table>
<thead>
<tr>
<th>Institution</th>
<th>Country</th>
<th>NP</th>
<th>NP (%)</th>
<th>NPC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>National University of Singapore</td>
<td>Singapore</td>
<td>13</td>
<td>4.8</td>
<td>100</td>
</tr>
<tr>
<td>Vermont Law School</td>
<td>USA</td>
<td>11</td>
<td>4.1</td>
<td>26.2</td>
</tr>
<tr>
<td>University of Oxford</td>
<td>England</td>
<td>10</td>
<td>3.7</td>
<td>11.2</td>
</tr>
<tr>
<td>International Institute for Applied Systems Analysis</td>
<td>Austria</td>
<td>8</td>
<td>3.0</td>
<td>57.1</td>
</tr>
<tr>
<td>University of Ulster</td>
<td>Northern Ireland</td>
<td>7</td>
<td>2.6</td>
<td>77.8</td>
</tr>
<tr>
<td>University of Birmingham</td>
<td>England</td>
<td>6</td>
<td>2.2</td>
<td>6.7</td>
</tr>
<tr>
<td>University College Dublin</td>
<td>Ireland</td>
<td>6</td>
<td>2.2</td>
<td>60.0</td>
</tr>
<tr>
<td>University of East Anglia</td>
<td>England</td>
<td>5</td>
<td>1.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Northumbria University</td>
<td>England</td>
<td>5</td>
<td>1.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Sheffield Hallam University</td>
<td>England</td>
<td>5</td>
<td>1.9</td>
<td>5.6</td>
</tr>
<tr>
<td>University College London</td>
<td>England</td>
<td>5</td>
<td>1.9</td>
<td>5.6</td>
</tr>
<tr>
<td>University of Edinburgh</td>
<td>Scotland</td>
<td>5</td>
<td>1.9</td>
<td>38.7</td>
</tr>
<tr>
<td>London University Imperial College</td>
<td>England</td>
<td>5</td>
<td>1.9</td>
<td>5.6</td>
</tr>
<tr>
<td>University of Otago</td>
<td>New Zealand</td>
<td>5</td>
<td>1.9</td>
<td>83.3</td>
</tr>
<tr>
<td>World Bank</td>
<td>USA</td>
<td>5</td>
<td>1.9</td>
<td>11.9</td>
</tr>
</tbody>
</table>

Notes: NP: number of publications; NP (%): percentage of publications considered; NPC (%): percentage of publications considered from one country in which the institution located.

3.5 Citation analysis

The citation count was obtained from Web of Science on 21 January 2014. The 269 publications have got 1398 non-self-citations, and the average non-self-citations per article was 5.2. Of these 269 publications, 79 have not been cited, accounting for 29.4%, and 40 have been cited just once, accounting for 14.9%. The overall impact of energy poverty publications was much less than other energy topics, such as climate change. The most cited publication was ‘Excess winter mortality in Europe: a cross country analysis identifying key risk factors’, which was published in 2003 by *Journal of Epidemiology and Community Health* (with an IF of 3.9), written by Healy, J.D. from University College Dublin. This publication has been cited 117 times (non-self-citation) since its publication. Table 4 show the top five most cited publications, and England has contributed two of them.
Table 4  Most cited publications in energy poverty research

<table>
<thead>
<tr>
<th>Author and year</th>
<th>Country</th>
<th>Journal</th>
<th>C</th>
<th>CPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healy (2003)</td>
<td>Ireland</td>
<td><em>Journal of Epidemiology and Community Health</em></td>
<td>117</td>
<td>9.8</td>
</tr>
<tr>
<td>Pachauri et al. (2004)</td>
<td>Switzerland</td>
<td><em>World Development</em></td>
<td>50</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Notes: Country: the country of the first unit of the latest publication of the first author; C: number of non-self-citations received; CPY: citations per year.

3.6 Collaboration

The author address information of each publication from Web of Science was used to analyse the collaboration. One hundred seventy-six publications were finished by different authors, 97 were cooperated by different institutions and 38 were international collaborations. The apparent increasing trend in the number of all three kinds of collaboration publications can be noted in Figure 3. There were an average of 2.6 authors, 1.6 institutions and 1.3 countries in each energy poverty publication considered.

Figure 3  The trend collaboration publications among authors, institutions and countries (see online version for colours)

Figure 4 presents the main collaboration relationship among countries, which have collaborated more than once. As shown in Figure 4, USA was the centre of international collaboration network, and developing countries were not full participants. USA has built collaboration relationships in the field of energy poverty with 14 countries, and with eight of them it has collaborated more than once. In the main network of international collaboration, there were only two developing countries, India and Nigeria, and they did not play important roles.
3.7 Hot topics

In order to capture the hot topics and major research trends, this paper analysed the keyword information. There were 499 words listed, of which 424 (85.0%) were used only once and 41 (8.2%) just twice. The most frequently used keywords were ‘energy poverty’ (59 times) and ‘fuel poverty’ (50 times), which were also the keywords used in searching.

With regard to the synonymous terms, spelling variations and abbreviations of the keywords, this paper summarised five hot topics based on the results of frequency analysis of keywords and artificial selection: fuel access, energy efficiency, physical health, electrification and income poverty (see Table 5). Fuel access includes keywords about relying on traditional biomass and lacking of commercial fuel, which are common in developing countries, especially in rural areas. These keywords have been used 70 times; energy efficiency considers improving the efficiency of housing and equipment by changing the building materials, introducing new stoves and biofuels; these keywords have been mentioned 65 times. Physical health includes the content about health hazards caused by energy poverty, such as indoor air pollution and excess winter mortality. These kinds of keywords have been used 55 times. Electrification refers to getting and consuming electricity, and upgrading power grids to ensure the quality of electricity. These keywords have been mentioned 49 times. Income poverty relates to the low-income groups and the least-developed countries, which are energy poverty vulnerable groups. This is the fifth hot issue in the field of energy poverty.
Table 5  The top 5 hot topics of energy poverty research

<table>
<thead>
<tr>
<th>Hot topics</th>
<th>Frequency</th>
<th>Main content</th>
<th>Masterpiece</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel access</td>
<td>70</td>
<td>Consume traditional biomass and commercial fuel</td>
<td>Kaygusuz (2011)</td>
</tr>
<tr>
<td>Energy efficiency</td>
<td>65</td>
<td>Improve efficiency of housing and equipment</td>
<td>Hong et al. (2009)</td>
</tr>
<tr>
<td>Physical health</td>
<td>55</td>
<td>Health hazards caused by energy poverty</td>
<td>Rudge and Gilchrist (2005)</td>
</tr>
<tr>
<td>Electrification</td>
<td>49</td>
<td>get and consume electricity</td>
<td>Pereira et al. (2010)</td>
</tr>
<tr>
<td>Income poverty</td>
<td>42</td>
<td>Energy poverty vulnerability of low-income groups</td>
<td>Khandker et al. (2012)</td>
</tr>
</tbody>
</table>

The top five hot topics are not independent of each other. Figure 5 shows their relationship. Lacking of commercial fuel and electricity is common in low-income groups and the least developed countries; relying on traditional biomass may cause serious indoor air pollution, and low heating efficiency may cause excess winter mortality; both electrification and renewable energy are good ways to improve energy efficiency; consuming traditional biomass is an inefficient behaviour; getting access to electricity is an important way to improve energy efficiency.

Figure 5  The relationship among top 5 hot topics in energy poverty research (see online version for colours)
Figure 6 shows the research trends of these five hot topics. Among 269 energy poverty publications, the first publication which had a keyword included in the top five topics was published in 2002. The frequency of all five topics showed significantly increasing trends, specifically energy efficiency and electrification.

Among the keywords were 26 countries, which were used 47 times. From the perspective of development, eight were developed countries and 18 were developing countries. More developing countries have been the focus of energy poverty. From the perspective of continents, nine countries were in Europe and eight were in Asia. Compared with other continents, Europe and Asia have got more attention. From the perspective of countries, India and England were mentioned six and five times, respectively. They have been the most important energy poverty research objectives.

4 Assessment model on energy poverty

In order to measure the academic research level on energy poverty of the main countries, this paper selected six indicators: number of publications, number of non-self-citations received, most cited publications (in top 20), most productive authors (in top 20), most productive institutions (in top 20) and most productive journals (in top 20). The
sum of standard scores of these six indicators is used to represent the level of energy poverty academic research of each country, and the assessment model used is (Wei et al., 2013)

\[
T_j = \frac{x_{ij} - \bar{x}_j}{\sqrt{\sum_j (x_{ij} - \bar{x}_j)^2 / (M - 1)}}
\]

\[
T_i = \sum_j w_j T_j
\]

where \(T_{ij}\) is the standard score of country \(i\) for indicator \(j\); \(T_i\) represents the academic research level on energy poverty of country \(i\); \(x_{ij}\) is the initial value of indicator \(j\) of country \(i\); \(\bar{x}_j\) is the average value of indicator \(j\) of countries which are involved in ranking; \(M\) is the number of countries which are involved in ranking; \(w_{ij}\) is the weight of country \(i\) for indicator \(j\). This paper defines that every indicator has the same weight and \(\sum_j w_{ij} = 1\). This paper compared the top 25 countries, which were ranked according to the number of publications. Figure 7 displays the standard scores of six indicators of 25 countries, and Figure 8 shows the total score which is on behalf of the academic research level on energy poverty.

**Figure 7** Standard scores of 6 indicators of 25 countries (see online version for colours)

The 25 countries were divided into four echelons based on the total scores (see Figure 8). Only England was in the first echelon. England’s score of the most productive authors equalled that of USA; other five standard scores were more than that of other countries,
and the total score of England was 25.0, about two times that of USA. The reasons why England came in first may be that energy poverty research started early in England, and energy poverty has become a formal government legislation in England. The second echelon included USA, with a total score of 9.6. Energy poverty research of USA started from 2005 with a rapid growth rate, and USA has built a widely international collaboration network. The third echelon had eight countries. These countries performed well in at least two indicators, and all of them were developed countries. Fifteen countries belonged to the fourth echelon. The differences among total scores of those countries were insignificant. China’s academic research level on energy poverty was ranked 12th. Although it was ahead of other developing countries, China did not perform very well in any indicator.

Figure 8  The academic research level on energy poverty of 25 countries (see online version for colours)

5 Conclusions

Energy poverty has a detrimental influence upon health, restricts the development of education, aggravates local ecological damage and can be the cause of a vicious cycle of a lack of economic and social development. This paper analysed the energy poverty literature of SCI-E and SSCI databases by bibliometric applications to identify publication patterns, the range of hot topics and the academic research level on energy poverty.

Energy poverty research just stepped into the initial development stage. The number of energy poverty publications was 269 of SCI-E and SSCI databases from 1981 to 2013. The most productive journal was Energy Policy, the most productive institution was the
National University of Singapore, the most productive author was Sovacool, who worked at the National University of Singapore and the Vermont Law School in succession, and the most cited publication was written by Healy, who was from Ireland.

Fuel access, energy efficiency, physical health, electrification and income poverty were all hot research topics in the field of energy poverty. Developing countries’ energy poverty has got more attention. Europe and Asia and India and England were the most important research objectives, from the perspective of continents and countries, respectively.

The energy poverty academic research capability of developed countries was much stronger than that of developing countries according to the results of the assessment model. The top ten most productive authors, institutions and journals all came from developed countries. India was the only developing country which published more than ten energy poverty articles. England’s energy poverty research capability ranked first, and developing countries were all in the fourth echelon.

Reference


DEFRA (Department of the Environment, Food and Rural Affairs) and DTI (Department of Trade and Industry) (2001) The UK fuel poverty strategy, DEFRA and DTI, London.


A bibliometric analysis of energy poverty research


Long-term trends in non-renewable resource commodity prices: fresh evidence in the presence of structural breaks

Aviral Kumar Tiwari

Faculty of Management,
IBS Hyderabad,
IFHE University Hyderabad,
Hyderabad 501203, Telangana, India
Email: aviral.eco@gmail.com
Email: aktiwari@ibsindia.org

Abstract: This study examines temporal properties of real price series of 11 natural resources covering the period 1870–1990 and utilising a recently developed unit root test proposed by Narayan and Poop and a stationarity test developed by Carrion-i-Silvestre and Sansó. We also employed a stationarity test developed by Lima and Neri. Results obtained from Lima and Neri’s study provide evidence to reject the null hypothesis of stationarity for aluminium, coal, copper, petroleum, silver and tin, indicating that these series are \( I(1) \), while results from the Carrion-i-Silvestre and Sansó proposed test provide evidence to reject the null hypothesis of stationarity for aluminium, iron, tin and zinc, indicating that these series are \( I(1) \). The results obtained from the Narayan and Poop unit root test provide significant evidence of stationarity for aluminium, copper, lead, silver, gas and zinc, indicating that only these variables are \( I(0) \).

Keywords: non-renewable resource commodity prices; structural breaks.


Biographical notes: Aviral Kumar Tiwari is Assistant Professor of Applied Economics in the Management Department at the IBS Hyderabad, IFHE University Hyderabad, India. His research interest is energy and environmental economics, macro-economics and applied econometrics. He has published in major journals of international repute such as Energy Economics, Energy Policy, Renewable & Sustainable Energy Reviews, Computational Economics, Empirical Economics, Economic Modelling, Quality & Quantity, Central European Journal of Economic Modelling and Econometrics etc. He serves as Editor-in-Chief of the Journal of Economic and Financial Modelling.

1 Introduction

Lots of studies have examined the behaviour of the price series of non-renewable resources with varied motivation. Noteworthy studies include Barnett and Morse (1963),
A.K. Tiwari

Smith (1979), Slade (1982), Berck and Roberts (1996), Ahrens and Sharma (1997) and Lee et al. (2006). These studies analysed, in general, whether non-renewable resources are becoming scarcer as a reflection of swelling prices, but none of them could arrive at a solid conclusion. Narayan and Liu (2010) divided these studies on commodity prices into three groups. The first group of studies, which included Akgiray et al. (1991) and Urich (2000), examined the distributional properties of future prices. The second group of studies included Fama and French (1988), Christie-David et al. (2000) and Cai et al. (2001), which analysed the effect of business cycle and macroeconomic news releases on futures prices of precious metals. The third group of studies included those of Dhillon et al. (1997) and Xu and Fung (2005) which examined the relationship between the cash market and the futures market or the metal futures trading in multiple markets.

Understanding the nature of resource prices’ time paths is important for several reasons. Firstly, knowledge of the correct time series behaviour of natural resource prices can be crucial to distinguish between theories that most accurately describe observed behaviour (Lee et al., 2006). Secondly, it is also important for proper econometric estimation such as co-integration. Thirdly, if a structural break is undetected or ignored, it may induce non-constancy in model parameters which will affect model interpretability (Mariscal and Powell, 2014). Fourthly, pretesting for unit roots with structural changes may help identify the most accurate forecasting model (as if a break goes undetected or is ignored, it may induce forecast failure), which is needed for good policymaking particularly for commodity-dependent countries. For example, in Latin America and the Caribbean region, some eight countries derive about a third of their fiscal revenues on average from non-renewable sources. And in a further set of countries, agricultural commodities are important foreign currency and tax earnings. Thus, knowledge of the persistence of prices, the likelihood that a structural break has occurred and how fast prices tend to revert to an equilibrium value serves as critical information for macroeconomic policymaking for such countries (Mariscal and Powell, 2014). Mariscal and Powell (2014) argued that the “identification of breaks and the development of a stable model for commodity prices are the important elements for the appropriate design of a stabilisation fund for commodity revenues, as implemented by Chile and other countries” (p.4). Lastly, we strongly believe the findings from our study have strong policy implications to arbitrageurs and speculators in the commodity trading market. As Xu and Fung (2005) show, arbitrageurs and speculators keenly follow metal (commodity) prices globally, and because metal commodities are characterised by a standard quality it enables arbitrage in cross-market futures trading. Thus, we make a direct contribution to the functioning of market participants in the commodities market.

More recently, interest of researchers has shifted to empirically examining whether the path of non-renewable natural resource prices is trend or difference stationary. For example, Slade (1982) examined the integration property of commodity prices of eight commodities using a Hotelling-type linear trend model in the spirit of a random walk type difference stationary model. Slade (1982) found that prices of seven of the eight commodities were characterised by a random walk. Similarly, Berck and Roberts (1996) examined the unit root properties of nine commodity prices, using annual data for the period from 1940 to 1991 and the Lagrange Multiplier (LM) test proposed by Schmidt and Phillips (1992) and the conventional Augmented Dickey–Fuller (ADF) test. Berck and Roberts (1996) found evidence of stationarity only for the price of silver. Ahrens and
Sharma (1997) using annual data on 11 commodity price series, ranging from 1870 to 1990, concluded that six of these series are stationary around a deterministic trend, while the remaining five display stochastic trends, implying a unit root. Using the same data set, and applying the LM unit root test and allowing for two endogenously determined structural breaks with and without a quadratic trend, Lee et al. (2006) documented that natural resource prices are stationary around deterministic trends. Recently, Sharma et al. (2009) examined the relationship between the number of structural breaks in the data and the nature of the resource price path, i.e. whether it is stationary or a random walk. Sharma et al. (2009) used Bai and Perron’s (1998) multiple structural break dating method and found that these series are, in many cases, stationary1 and subject to a number of structural breaks. Hence, Sharma et al. (2009) concluded that a deterministic model of resource prices may well be appropriate. Most recently, Narayan and Liu (2010) examined whether shocks to ten commodity prices (gold, silver, platinum, copper, aluminium, iron ore, lead, nickel, tin and zinc) are persistent or transitory by using daily data and applying two recently developed unit root tests, namely the Narayan and Popp (NP) (2010) test and the Liu–Narayan (LN) test (Liu and Narayan, 2010) test that allow for two structural breaks in the data series. Narayan and Liu (2010) found that shocks to gold, silver, platinum, aluminium and copper are persistent.

In our work we utilised the same data used in Ahrens and Sharma (1997), Lee et al. (2006) and Sharma et al. (2009), and we made an attempt to provide new results by using a more powerful and recently developed unit root test as well as a stationarity test which incorporates up to two endogenously determined structural breaks in the data, such as the Narayan and Popp (2010) and Carrion-i-Silvestre and Sansó (2007) tests respectively. It is well known that ignoring structural change in unit root tests will lead to a bias against rejecting the unit root null hypothesis when it should, in fact, be rejected. Further, we also used several other recently developed more powerful stationarity tests (which do not incorporate structural breaks explicitly) together with ‘classical’ unit root tests to compare the results. The other powerful tests of a stationarity null hypothesis are robust to unit root alternative, alternatives to structural changes in the mean and alternatives with unconditional heteroskedasticity, and have good power in detecting changes in higher moments of the unconditional distribution.

The remainder of the paper is structured as follows. Section 2 provides data source and describes our empirical methodologies. Section 3 presents our empirical results. Section 4 concludes.

2 Data and methodology

2.1 Data

The data of the non-renewable natural resources that we used in this paper have previously been used by Ahrens and Sharma (1997), Lee et al. (2006) and Sharma et al. (2009). A complete description of data construction and the various sources used can be found in Ahrens and Sharma (1997, pp.66–67). A summary of the data periods covered by each series is presented in Table 1. As can be seen from Table 1, data periods somewhat vary.
Table 1  Data summary

<table>
<thead>
<tr>
<th>Non-renewable resource</th>
<th>Abbreviation</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium</td>
<td>AL</td>
<td>1895–1984</td>
</tr>
<tr>
<td>Coal</td>
<td>CL</td>
<td>1870–1990</td>
</tr>
<tr>
<td>Copper</td>
<td>CP</td>
<td>1870–1990</td>
</tr>
<tr>
<td>Iron</td>
<td>IR</td>
<td>1870–1973</td>
</tr>
<tr>
<td>Lead</td>
<td>LE</td>
<td>1870–1990</td>
</tr>
<tr>
<td>Nickel</td>
<td>NI</td>
<td>1913–1990</td>
</tr>
<tr>
<td>Petroleum</td>
<td>PT</td>
<td>1870–1989</td>
</tr>
<tr>
<td>Silver</td>
<td>SI</td>
<td>1870–1990</td>
</tr>
<tr>
<td>Tin</td>
<td>TI</td>
<td>1885–1990</td>
</tr>
<tr>
<td>Gas</td>
<td>GA</td>
<td>1919–1990</td>
</tr>
<tr>
<td>Zinc</td>
<td>ZI</td>
<td>1870–1990</td>
</tr>
</tbody>
</table>

2.2 Methodology

In the estimation methodology, we first used unit root tests those are utilised more often in the economic literature such as the ADF test (Dickey and Fuller, 1981) and the Phillips–Perron (PP) test (Phillips and Perron, 1988) together with other unit root tests. We call them ‘classical’ unit root tests. Secondly, we used recently developed more powerful tests of stationarity, which does not take into account the existence of structural breaks in the null or alternative hypothesis. Thirdly, we used unit root tests that incorporate the structural breaks in either the null hypothesis or the alternative hypothesis. We briefly discuss in the following sections the unit root tests used in the study.

2.2.1 ‘Classical’ unit root tests

In the first step, we used the most popular ADF (Dickey and Fuller, 1981) and the PP unit root tests (Phillips and Perron, 1988). The augmented form of the DF test is used when there is the problem of serial correlation, and to choose appropriate lag length the Schwarz Information Criterion (SIC) is preferred. Since the PP test has advancements over the ADF test in the sense that the ADF test uses a parametric auto-regression to approximate the ARMA structure of the errors in the test regression, while the PP test corrects for any serial correlation and heteroskedasticity in the errors in the non-parametric framework. Therefore, it is also used for analysis. In the PP test to select appropriate lag length, we have adopted the Newey–West method using the Bartlet kernel method. Further, we also use some other popular unit root tests such as the Elliott–Rothenberg–Stock test (Elliott et al., 1996) and the Ng–Perron test (Ng and Perron, 2001).

2.2.2 Powerful stationarity tests

Kwiatkowski et al. (1992) proposed a test (KPSS test) (generally known as a KPSS test statistic) for the first-order (level) stationarity, which is defined as follows:

\[
KPSS = \frac{1}{(wT)} \left( \sum_{i=1}^{T} \left( \sum_{j=1}^{T} (Y_i - \bar{Y}) \right)^2 \right)
\]
where $\bar{y}_i$ is the sample mean of $\{y_i\}_{i=1}^T$ and $\hat{w}^2$ is a non-parametric consistent estimator of the long-run variance. Under the null hypothesis of level stationarity, the KPSS test statistic can also be represented as

$$KPSS \Rightarrow \int_0^1 \kappa(\alpha)^2 d\alpha$$

(2)

where $\kappa(\alpha) = W(\alpha) - W(\alpha)(1)$ is the standard Brownian bridge.

Recently, de Jong et al. (2007) proposed a robust version of the KPSS test (we call it IKPSS test statistic) based on the following empirical process:

$$I_r(r) := \frac{1}{\sigma \sqrt{T}} \sum_{i=1}^{[T]} \text{sign}(y_i - m_T)$$

(3)

where $m_T$ is the sample median of $\{y_i\}_{i=1}^T$ and $\hat{\sigma}^2$ is a non-parametric consistent estimator of the long-run variance

$$\sigma^2 = \lim_{T \to \infty} E\left[\frac{1}{\sqrt{T}} \sum_{i=1}^{[T]} \text{sign}(y_i - m_T)^2\right]$$

and

$$\text{sign}(x) := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

The fluctuation of the empirical process $I_r(r)$ is measured by $h(I_r(r))$ where $h(\cdot)$ is the Cramér–von Mises metric. Thus, the IKPSS test statistic can be expressed as

$$\text{IKPSS} := \frac{1}{(\hat{\sigma}T)^2} \sum_{i=1}^T \left(\sum_{j=1}^d \text{sign}(y_{i,j} - m_T)^2\right)^2$$

(4)

de Jong et al. (2007) show that under the null hypothesis of level stationarity, IKPSS test statistic has the same limiting distribution as the KPSS test statistic, i.e.

$$\text{IKPSS} \Rightarrow \int_0^1 \kappa(\alpha)^2 d\alpha.$$  

We should note that when the alternative hypothesis is unit root, the IKPSS has the correct size under the presence of fat-tailed errors, while the KPSS test does not and when the tails are thin the IKPSS test has a lower power than the KPSS test. However, when the aforementioned traditional stationarity tests are applied to test stationarity, it is difficult to detect alternatives with unconditional volatility (distribution scale) that changes over time. To overcome this issue, Xiao and Lima (2007) proposed a test (we call it XL) for second-order (covariance) stationarity based on the following standardised bivariate empirical process:

$$Z_r(r) := \frac{1}{\sqrt{T}} \hat{\Omega}^{-1/2} \sum_{i=1}^T \left(\tilde{y}_i, v_i\right)$$

(5)
where $\bar{y}_t = y_t - \frac{1}{T} \sum_{j=1}^{T} y_j$ is the demeaned data, $v_t := \bar{y}_t^2 - \sigma_t^2$, $\sigma_t^2 := \frac{1}{T} \sum_{t=1}^{T} \bar{y}_t^2$ and $\hat{\Omega}^{1/2}$ is the inverse of the Cholesky decomposition of $\hat{\Omega}$, a non-parametric consistent estimator of the long-run variance

$$\Omega^2 = \lim_{T \to \infty} E \left[ \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \bar{y}_t \right)^2 \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} v_t \right)^2 \right]$$

Then, Xiao and Lima (2007) applied the Kolmogorov metric to measure the fluctuation of the empirical process $Z_T(r)$ and defined their test statistic as follows:

$$XL = \max_{1 \leq k \leq T} \left\| \frac{1}{\sqrt{T}} \hat{\Omega}^{1/2} \sum_{t=1}^{T} \bar{y}_t \right\| \quad (6)$$

Under the null hypothesis of covariance stationarity, the test statistic can be expressed as

$$XL \Rightarrow \sup_{0 \leq r \leq 1} \left\| W_1(r) - rW_1(1) \right\| \quad (7)$$

where $(W_1(r) - rW_1(1), W_2(r) - rW_2(1))'$ is the two-dimensional standardised Brownian bridge. The critical values can be found in Xiao and Lima (2007). Unlike the KPSS or the IKPSS, the XL test has power not only against the alternative hypothesis of distribution location varying in time, but also against the alternative hypothesis of the distribution scale (unconditional volatility) varying in time. However, all of the aforementioned tests have power close to size against the alternative hypothesis of time-varying kurtosis. As Busetti and Harvey (2007) discuss, the distribution of a random variable may present changes over time that does not impact the level or the variance.

For instance, maybe the asymmetry or fatness of the tail is time varying. This is particularly important in analysing financial time series. To exemplify this point, consider how changes in lower tail quantiles may impact decisions of a risk manager or a regulatory agency.

To overcome the problem of the aforementioned unit root test, Lima and Neri (2013) propose a new test (which we call the LN test) for the null hypothesis of strict stationarity generalising the IKPSS test in terms of using the sign of the data minus the sample quantiles, whereas the IKPSS test uses the sign of the data minus the sample median only. Not only does the LN test have power against unit root alternative, alternatives to structural changes in the mean and alternatives with unconditional heteroskedasticity, but it also has good power in detecting changes in higher moments of the unconditional distribution unlike the KPSS, IKPSS and XL tests. The estimation procedure of the LN test can be explained as follows.

Let $\{y_t\}_{t=1}^{T}$ be the data and, for $\tau \in [0,1]$, define

$$b(\tau) := \arg \max_{b \in \mathbb{R}} \sum_{t=1}^{T} \rho_\tau(y_t - b) \quad (8)$$

where $\rho_\tau(u) = (1_{|u| < \tau} - \tau)u$. 

```
Therefore, \( b(\tau) \) is simply the \( \tau \)th sample unconditional quantile of \( \{y_t\}_{t=1}^T \). Notice that \( \rho_\tau \) is not everywhere differentiable, but since it is convex, we can still compute the sub-gradient. The sub-gradient plays the same role in quantile estimation as the score function in maximum likelihood estimation. The sub-gradient of \( \rho_\tau \) is given by

\[
\psi(u) = 1_{u < 0} - \tau.
\]

We now define the empirical process

\[
S_y(r, \tau) := \frac{1}{\hat{\pi}(\tau) \sqrt{T}} \sum_{t=r}^{\lfloor T \rfloor} \psi_r(y_t - b(\tau))
\]

where \( r \in [0,1] \) and \( \hat{\pi}(\tau) \) is a non-parametric consistent estimator of

\[
\pi(\tau) := \lim_{T \to \infty} E \left[ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \psi_r(y_t - b_0(\tau)) \right]^2
\]

where \( b_0(\tau) \) is the population \( \tau \)th unconditional quantile of \( \{y_t\}_{t=1}^T \).

By using the Kolmogorov–Smirnov metric to measure the fluctuation of \( S_y(r, \tau) \) across various quantiles \( \tau \in \Gamma_w = [w, 1 - w] \), for some \( w \in (0,1/2) \), the LN test statistic for strict stationarity can be expressed as follows:

\[
LN = \max_{r, w} \max_{1 \leq k \leq T} \frac{1}{\hat{\pi}(\tau) \sqrt{T}} \left| \sum_{t=r}^{\lfloor T \rfloor} \psi_r(y_t - b(\tau)) - \frac{k}{T} \sum_{t=1}^{t} \psi_r(y_t - b(\tau)) \right|
\]

\( \hat{\pi}(\tau) \) can be computed as the HAC estimator

\[
\hat{\pi}(\tau) := \frac{1}{T} \sum_{j, i} K \left( \frac{i - j}{q_T} \right) \psi_r(y_i - b(\tau)) \psi_r(y_j - b(\tau))
\]

where \( K \) is a kernel function.

### 2.2.3 Unit root with structural breaks

To overcome the problem of ‘classical’ and powerful stationarity tests in detecting the structural breaks and incorporating break in the null or alternative hypothesis, several tests are proposed in the literature, such as the ZA and LP tests. However, Lee and Strazicich (2003, 2004) observed that ZA and LP tests identify the break point one period prior to the true break point (i.e. \( T_{Bt-1} \) rather than \( T_{Bt} \)) and the bias in estimating the persistence parameter is maximised and spurious rejections are the greatest. To overcome this limitation, Lee and Strazicich (2003) proposed the minimum LM-based unit root test with two structural breaks. Lee and Strazicich (2004) proposed another LM unit root test with one break. Popp (2008) observed that the root of the problem of spurious rejections is that the parameters of the test regression have different interpretations under the null and the alternative hypotheses, which are crucial since the parameters have implications for the selection of the structural break. Narayan and Popp (2010) dealt with this problem (following Schmidt and Phillips 1992) by developing an ADF-type test for the case of Innovational Outlier (IO), where the data generating process is formulated as an
unobserved component model. Narayan and Popp (2010) claim that in the new test “Critical Values (hereafter CVs) of the test, assuming unknown break dates, converge with increasing sample size to the CVs when break points are known”. Further, Narayan and Popp (2010) claim that the “critical values (CVs) of the test, assuming unknown break dates, converge with increasing sample size to the CVs when break points are known”. Therefore, it identifies the break point more accurately than the earlier tests and Narayan and Popp (2010) verified that the rejection frequency is relatively less in their test. Therefore, we used the NP test as well in our analysis. Narayan and Popp (2010) considered following types of DGP of a time series, $y_t$, which has two components, a deterministic component ($d_t$) and a stochastic component ($u_t$):

$$y_t = d_t + u_t$$  \hspace{1cm} (11)

$$u_t = \rho u_{t-1} + \varepsilon_t$$  \hspace{1cm} (12)

and

$$\varepsilon_t = \Psi(L)e_t = A'(L)^{-1}B(L)e_t$$  \hspace{1cm} (13)

where $e_t$ is a white noise process, such that $e_t \sim N(IID)(0, \sigma^2)$. Narayan and Popp (2010) assumed that the roots of the lag polynomials $A'(L)$ and $B(L)$, which are of order of $p$ and $q$, respectively, lie outside the unit circle. With this assumption, Narayan and Popp (2010) considered two different specifications for trending data: one allows for two breaks in level (denoted as model 1, i.e. M1) and the other allows for two breaks in level as well as the slope (denoted as model 2, i.e. M2). The specification of both models differs in terms of the definition of the deterministic component, $d_t$:

$$d_t^{M1} = \alpha + \beta t + \Psi(L)\left(\theta_1 DU_{t,i}^{U} + \theta_2 DU_{t,j}^{U}\right)$$  \hspace{1cm} (14)

$$d_t^{M2} = \alpha + \beta t + \Psi(L)\left(\theta_1 DU_{t,i}^{U} + \theta_2 DU_{t,j}^{U} + \gamma_1 DT_{t,i}^{T} + \gamma_2 DT_{t,j}^{T}\right)$$  \hspace{1cm} (15)

with

$$DU_{t,i}^{U} = 1(t > T_{t,i}^{U}) , DT_{t,i}^{T} = 1(t > T_{t,i}^{T})(t - DT_{t,i}^{T}) , i = 1, 2$$  \hspace{1cm} (16)

where $T_{t,i}^{U}, i = 1, 2$, denotes the true break dates and $\theta_1$ and $\gamma_1$ indicate the magnitude of the level and slope breaks, respectively. The inclusion of $\Psi(L)$ in equations (14) and (15) allows the breaks to occur slowly over time, i.e. it assumes that the series responds to shocks to the trend function the way it reacts to shocks to the innovation process $e_t$ (Vogelsang and Perron, 1998) and so this process is known as the IO model. Narayan and Popp (2010) derived the IO-type test regressions to test for the unit root hypothesis for M1 and M2 by merging the structural models (11)–(15). The test regressions can be derived from the corresponding structural models in reduced form as follows:

$$y_t^{M1} = \rho y_{t-1} + \alpha_1 + \beta'_t t + \theta_1 D(T_{t}^*)_{i,j} + \theta_2 D(T_{t}^*)_{2,j}$$  

$$+ \delta_1 DU_{t,i}^{U} + \delta_2 DU_{t,j}^{U} + \sum_{j=1}^{i} \beta_j D\Psi_{t,j} + \varepsilon_t$$  \hspace{1cm} (17)
where equations (17) and (18) are IO-type test regression for M1 and M2, respectively. In order to test the unit root null hypothesis of $\rho = 1$ against the alternative hypothesis of $\rho < 1$, we use the $t$-statistics of $\hat{\rho}$, denoted $t_\rho$, in equations (17) and (18). Since it is assumed that true break dates are unknown, $T^*_b$, in equations (17) and (18) has to be substituted by their estimates $\hat{T}^*_b$, $i = 1, 2$, in order to conduct the unit root test. The break dates can be selected simultaneously following a grid search procedure or a sequential procedure comparable to Kapetanios (2005). Narayan and Poop (2010) preferred the sequential procedure because it is far less computationally demanding; therefore, we have also followed the sequential procedure.

The first step in this case is the search for a single break, according to the maximum absolute $t$-value of the break dummy coefficient $\theta_1$ for M1 and $\kappa_1$ to the M2. Thereafter, we impose the restriction $\theta_2 = \delta_2 = 0$ for M1 and $\kappa_2 = \delta = \gamma = 0$ for M2, and hence we have

\[
T_{b,1}^* = \arg\max_{\hat{T}_b} \{ f_{\theta_1}(\hat{T}_b) \}, \quad \text{for M1}
\]
\[
T_{b,2}^* = \arg\max_{\hat{T}_b} \{ f_{\theta_2}(\hat{T}_b) \}, \quad \text{for M2}
\]

Therefore, in the first step, the test procedure reduces to the case described in Popp (2008). Imposing the first break $\hat{T}_{b,1}$ in the test regression, we estimate the second break date $\hat{T}_{b,2}$. Again, we maximise the absolute $t$-value, this time $\theta_2$ for M1 and $\kappa_2$ for M2. Hence, we have

\[
T_{b,2}^* = \arg\max_{\hat{T}_b} \{ f_{\theta_2}(\hat{T}_{b,1}, \hat{T}_{b,2}) \}, \quad \text{for M1}
\]
\[
T_{b,2}^* = \arg\max_{\hat{T}_b} \{ f_{\kappa_2}(\hat{T}_{b,1}, \hat{T}_{b,2}) \}, \quad \text{for M2}
\]

Now we will give a brief explanation of Carrion-i-Silvestre and Sansó’s (2007) proposed stationarity tests with two structural breaks. Carrion-i-Silvestre and Sansó (2007) proposed KPSS-based tests of stationarity (Kwiatkowski et al., 1992) with two structural breaks for seven different models. The standard KPSS test is based on

\[
y_t = f(t, T_{b1}, T_{b2}) + \epsilon_t + \eta_t
\]
\[
r_t = r_{t-1} + \epsilon_t
\]
where $e_t \sim \text{iid}\{0, \sigma_e^2\}$ and \{\$\} is assumed to satisfy the strong-mixing regularity conditions of Phillips and Perron (1988). Under the null hypothesis of stationarity $\sigma_e^2$ must be zero; otherwise, the stochastic process is I(1). $f(t, T_{B1}, T_{B2})$ in (21) denotes the deterministic specification that is assumed for the time series. In order to take into account the presence of the structural breaks, seven specifications presented in Table A1 incorporate dummy variables, which are defined as $DU_{ji} = \mathbb{1}(t > T_{Bj})$ if $t > T_{Bj}$ and 0 otherwise, with $T_{Bj} = \lambda_j T$, $\lambda_j \in (0, 1)$, $i = 1, 2$, denoting the date of the structural breaks. Carrion-i-Silvestre and Sansó (2007) proposed a pseudo-LM test, which is given as

$$
\hat{\eta}_j = \hat{\sigma}^2 T^{-2} \sum_{i=1}^{T} S_i^2
$$

where $j = \{\text{AA, AN, BB, CC, AB-BA, AC-CA, BC-CB}\}$ (see Table A1 for these models), where $S_i = \sum_{j=1}^{T} \hat{u}_j, S_0 = 0$ with $\hat{u}_j$ being the OLS estimated residuals of the regression of $y_t$ on one of the deterministic specifications in Table A1. Kwiatkowski et al. (1992) estimate the long-run variance from

$$
\hat{\sigma}^2 = T^{-1} \sum_{i=1}^{T} \hat{e}_i^2 + 2T^{-1} \sum_{l=1}^{L} w(s, l) \sum_{i=r+1}^{T} \hat{e}_i \hat{e}_{i-s}
$$

where $w(s, l)$ denotes the spectral window, i.e. either the Bartlett or the quadratic spectral windows. In the literature, there are varied suggestions for using a type of spectral window as it affects the results (for details, see Carrion-i-Silvestre and Sansó 2006a). In this study, therefore, we also use a boundary rule for the estimation of the long-run variance. We computed the pre-whitened Heteroskedasticity and Autocorrelation Consistent (HAC) estimator for the long-run variance as follows. In the first stage, an AR model for the residuals \{\$\} is estimated:

$$
\hat{e}_i = v_i \hat{e}_{i-1} + \ldots + v_p \hat{e}_{i-p} + \phi_i
$$

First, we estimate equation (24); then we obtained the long-run variance of the estimated residuals, denoted as $\hat{\sigma}^2_{\hat{\sigma}}$, through the application of an HAC estimator to control for the presence of heteroskedasticity. Finally, the estimated long-run variance is

$$
\hat{\sigma}^2 = \frac{\hat{\sigma}^2_{\hat{\sigma}}}{\hat{v}(1)^2}
$$

where $\hat{v}(1)$ denotes the autoregressive polynomial $\hat{v}(L) = 1 - \hat{v}_1 L - \ldots - \hat{v}_L L^L$ estimated in equation (14) evaluated at $L = 1$. Following Sul et al. (2005), we used

$$
\hat{\sigma}^2 = \min \left\{ k \hat{\sigma}^2 T, \frac{\hat{\sigma}^2_{\hat{\sigma}}}{\hat{v}(1)^2} \right\}
$$
boundary condition rule to obtain the long-run variance estimate, where $\kappa > 0$ is a constant to be determined below. Sul et al. (2005) used $\kappa = 1$, although it is suggested that other values of $\kappa$ might be suitable to improve the power of the test.

Now we will explain the estimation of the break points in Carrion-i-Silvestre and Sansó (2007). Carrion-i-Silvestre and Sansó (2007) followed the recommendation of Carrion-i-Silvestre and Sansó (2006a) to use the minimisation of the sequence of the Sum of Squared Residuals (SSR) to estimate the date of the break proposed by Kurozumi (2002). The procedure chooses the dates of the breaks from the argument that minimises the sequence of $SSR(T_{b1}, T_{b2})$, where SSR is obtained from the regression of $y_t = f(t, T_{b1}, T_{b2}) + e_t$ and $f(t, T_{b1}, T_{b2})$ denotes one of the deterministic components in Table 1. Thus, the break points are estimated as

$$\left(\hat{T}_{b1}, \hat{T}_{b2}\right) = \arg \min_{\Lambda} SSR(T_{b1}, T_{b2})$$  \hspace{1cm} (27)

where $\Lambda$ denotes a closed subset of the interval $(0,1)^2$. In order to minimise the loss of information, we define it as $\Lambda = [2/T, (T – 1)/T]^2$. Bai (1994, 1997) and Bai and Perron (1998), respectively, show the $T$-consistency of the estimation of the break fraction parameter when it is estimated using this criteria for the case of one structural break and for the case of multiple breaks for both trending and non-trending regressors. Carrion-i-Silvestre and Sansó (2006b) extended the same for models that include stochastic regressors. In the present case, we utilised the CC model for our analysis in order to compare our results of the KPSS test with NP (2010) test:

$$\left(\hat{T}_{b1}, \hat{T}_{b2}\right) = \arg \min_{\Lambda}. $$

### 3 Empirical results

To understand the nature of the time series utilised, we analysed the descriptive statistics and presented results in Table 2.

One can observe from the significant Jarque–Bera test statistics reported in Table 2 that all the series exhibit the non-normal property. The high standard deviation value shows that all the series are very volatile. Skewness statistics show that all variables are positively skewed and kurtosis statistics show that all series have high peak. Moreover, the Ljung–Box statistic suggests the presence of serial correlation in all the series and the Autoregressive Conditional Heteroskedasticity–Lagrange Multiplier (ARCH–LM) statistic indicates that ARCH effect is likely to be found in all the series. This nature of the series shows that the all series have a fat-tailed distribution; thus, as indicated in the ‘Methodology’ section, the more powerful stationarity tests are better to reach a conclusion. The results of the ‘classical’ unit root tests and powerful stationarity tests are presented in Table 3.
Table 2

Descriptive statistics of the variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Std. dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Jarque-Bera</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium</td>
<td>53.62088</td>
<td>39.31656</td>
<td>233.5119</td>
<td>18.70722</td>
<td>42.09166</td>
<td>2.018489</td>
<td>7.472972</td>
<td>136.1425</td>
<td>90</td>
</tr>
<tr>
<td>Coal</td>
<td>5.45603</td>
<td>4.8</td>
<td>12.59863</td>
<td>3.4</td>
<td>2.070406</td>
<td>1.65671</td>
<td>5.372627</td>
<td>83.73257</td>
<td>121</td>
</tr>
<tr>
<td>Copper</td>
<td>36.92562</td>
<td>36.4</td>
<td>70.7</td>
<td>16.6</td>
<td>11.4296</td>
<td>0.482581</td>
<td>2.632411</td>
<td>5.37774</td>
<td>121</td>
</tr>
<tr>
<td>Gas</td>
<td>23.38945</td>
<td>16</td>
<td>82.40077</td>
<td>7.8</td>
<td>19.60321</td>
<td>1.77978</td>
<td>5.040188</td>
<td>50.49849</td>
<td>72</td>
</tr>
<tr>
<td>Iron</td>
<td>101.374</td>
<td>99.5</td>
<td>181.4</td>
<td>63.3</td>
<td>21.81045</td>
<td>0.650253</td>
<td>3.550157</td>
<td>8.640613</td>
<td>104</td>
</tr>
<tr>
<td>Lead</td>
<td>13.69422</td>
<td>13.5</td>
<td>22.41691</td>
<td>6.072903</td>
<td>2.728663</td>
<td>0.138992</td>
<td>4.124304</td>
<td>6.762561</td>
<td>121</td>
</tr>
<tr>
<td>Nickel</td>
<td>83.98254</td>
<td>78.90501</td>
<td>188.1468</td>
<td>43.47826</td>
<td>25.73684</td>
<td>1.455826</td>
<td>6.364303</td>
<td>63.56</td>
<td>78</td>
</tr>
<tr>
<td>Petroleum</td>
<td>4.299339</td>
<td>3.77</td>
<td>14.71</td>
<td>0</td>
<td>2.267224</td>
<td>2.215089</td>
<td>8.577652</td>
<td>255.7974</td>
<td>121</td>
</tr>
<tr>
<td>Silver</td>
<td>199.2159</td>
<td>160.5256</td>
<td>1069.091</td>
<td>75.30169</td>
<td>122.7535</td>
<td>3.170907</td>
<td>21.90043</td>
<td>2003.785</td>
<td>121</td>
</tr>
<tr>
<td>Tin</td>
<td>119.5617</td>
<td>107.1692</td>
<td>299.2444</td>
<td>55.2083</td>
<td>50.09982</td>
<td>1.664103</td>
<td>5.986628</td>
<td>88.31965</td>
<td>106</td>
</tr>
<tr>
<td>Zinc</td>
<td>15.45314</td>
<td>15</td>
<td>40.2</td>
<td>1.4</td>
<td>3.931197</td>
<td>1.646701</td>
<td>16.23251</td>
<td>937.4772</td>
<td>121</td>
</tr>
</tbody>
</table>

Notes:  
*Rejection of the null hypothesis at the 1% level of significance.  
*Rejection of the null hypothesis at the 5% level of significance.  
*Rejection of the null hypothesis at the 10% level of significance.
Table 3: Results of ‘classical’ unit root tests and powerful stationarity tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADF(tau)</td>
<td>PP(Z-tau)</td>
<td>DF-GLS</td>
<td>P-test</td>
</tr>
<tr>
<td>Aluminium</td>
<td>–3.2301</td>
<td>–5.4094</td>
<td>–0.7886</td>
<td>88.9826</td>
</tr>
<tr>
<td>Coal</td>
<td>–3.7256</td>
<td>–2.8566</td>
<td>–3.1806</td>
<td>6.9699</td>
</tr>
<tr>
<td>Copper</td>
<td>–2.6835</td>
<td>–3.8451</td>
<td>–2.7676</td>
<td>7.2083</td>
</tr>
<tr>
<td>Gas</td>
<td>–1.699</td>
<td>–1.708</td>
<td>–1.6057</td>
<td>14.1809</td>
</tr>
<tr>
<td>Iron</td>
<td>–3.4901</td>
<td>–3.1332</td>
<td>–2.424</td>
<td>10.152</td>
</tr>
<tr>
<td>Lead</td>
<td>–3.6975</td>
<td>–4.7185</td>
<td>–3.7248</td>
<td>2.3653</td>
</tr>
<tr>
<td>Nickel</td>
<td>–2.8944</td>
<td>–3.8074</td>
<td>–1.9907</td>
<td>6.0272</td>
</tr>
<tr>
<td>Petroleum</td>
<td>–2.8267</td>
<td>–3.6133</td>
<td>–1.6944</td>
<td>20.4309</td>
</tr>
<tr>
<td>Silver</td>
<td>–2.1341</td>
<td>–4.3894</td>
<td>–2.0816</td>
<td>10.1856</td>
</tr>
<tr>
<td>Tin</td>
<td>–2.5432</td>
<td>–2.7497</td>
<td>–2.8812</td>
<td>5.7771</td>
</tr>
<tr>
<td>Zinc</td>
<td>–3.6442</td>
<td>–8.9874</td>
<td>–3.6571</td>
<td>2.5675</td>
</tr>
</tbody>
</table>

Notes: All computations in this table are done in R 2.15 (2012) and for normal unit root tests ‘urca’ package by Bernhard Pfaff is used and for powerful unit root tests we used codes developed by LN. Critical values related to ‘classical’ unit root tests are reported in Table A2.

*Rejection of the null hypothesis at the 1% level of significance.

*bRejection of the null hypothesis at the 5% level of significance.

^Rejection of the null hypothesis at the 10% level of significance.
Results for the "classical" unit root tests [particularly using the Ng and Perron (2001) test] show that coal, copper, lead, nickel and zinc are stationary variables and the remaining seven series are non-stationary. If we consider another unit root test result, we find no consistency in the outcome. Thus, we rely on the Ng and Perron (2001) test, as it is argued to be the more powerful test among the normal tests.

Table 4  Results of a unit root tests and a stationarity test analysis with structural breaks

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_{Bi}$</td>
<td>$T_{Bi}$</td>
</tr>
<tr>
<td></td>
<td>$M1(k)$</td>
<td>$T_{Bi}$</td>
</tr>
<tr>
<td>Aluminium</td>
<td>0.1339334b (0.0972) [1898,1907]</td>
<td>-3.296677 (1)</td>
</tr>
<tr>
<td>Coal</td>
<td>0.02012701a (0.0518) [1919,1973]</td>
<td>-3.572622 (4)</td>
</tr>
<tr>
<td>Copper</td>
<td>0.0318581 (0.0518) [1874,1923]</td>
<td>-4.108507 (0)</td>
</tr>
<tr>
<td>Iron</td>
<td>0.1423416a (0.0766) [1874,1923]</td>
<td>-1.806063 (5)</td>
</tr>
<tr>
<td>Lead</td>
<td>0.0318581 (0.0518) [1874,1922]</td>
<td>-5.492501a (1)</td>
</tr>
<tr>
<td>Nickel</td>
<td>0.0288608 (0.0608) [1969,1987]</td>
<td>-3.351967 (0)</td>
</tr>
<tr>
<td>Petroleum</td>
<td>0.0489086 (0.0766) [1878,1979]</td>
<td>-3.0576501 (2)</td>
</tr>
<tr>
<td>Silver</td>
<td>0.0792948 (0.0965) [1977,1980]</td>
<td>-3.484282 (2)</td>
</tr>
<tr>
<td>Tin</td>
<td>0.2877879a (0.0966) [1956,1976]</td>
<td>-3.4588463 (5)</td>
</tr>
<tr>
<td>Gas</td>
<td>0.03977345 (0.0766) [1972,1984]</td>
<td>-4.070926 (5)</td>
</tr>
<tr>
<td>Zinc</td>
<td>0.0184309 (0.0000) [1914,1916]</td>
<td>-5.309148a (5)</td>
</tr>
</tbody>
</table>

Notes: (1) CC denotes Carrion-i-Silvestre and Sansó (2007) proposed KPSS-type test, and M1 and M2 are the Narayan and Popp (2010) proposed test of Models 1 and 2. (2) 'k' denotes lag length. (3) $T_{Bi}$ denotes break dates for $i = 1, 2$. (4) For the NP (2010) test, critical values are obtained from Table 3. (5) For the KPSS test, the critical values at the 5% level of significance are obtained from the response surfaces in Table 2 and we report those in {}.

a Rejection of the null hypothesis at the 1% level of significance.
b Rejection of the null hypothesis at the 5% level of significance.
c Rejection of the null hypothesis at the 10% level of significance.

Source: Author’s calculation

Moreover, we can observe from the results obtained from the powerful stationarity tests that the null hypothesis of stationarity is rejected for aluminium, coal, copper and tin by all the stationarity tests used. For gas, only the KPSS test rejects the null hypothesis of stationarity. For petroleum, the null hypothesis of stationarity is rejected by the KPSS,
Long-term trends in non-renewable resource

IKPSS and LN tests, whereas for silver, only XL and LN tests reject the null hypothesis of stationarity. For zinc, only the IKPSS test rejects the null hypothesis of stationarity, but at a 10% significance level. Thus, our results show that iron, lead and nickel are strictly stationary and remaining eight series are strictly non-stationary. To further confirm these results, we used the NP (2010) unit root test and the Carrion-i-Silvestre and Sansó (2007) stationarity test, which incorporates structural break explicitly (although powerful results are robust to such cases, but do not take into account structural breaks explicitly) and we present the obtained results in Table 4.

It is evident from Table 4 that by using the Carrion-i-Silvestre and Sansó (2007) stationarity test, we could reject the null hypothesis of stationarity for aluminium, iron, tin and zinc, and for the rest of the series the null hypothesis of stationarity is not rejected. Now if we compare these results with those of the powerful unit root tests reported in Table 3, we find that incorporation of the structural break leads us to reject the null hypothesis of stationarity for iron; however, lead and nickel are still found to be stationary variables. Again, to see the robustness of these results, we used the NP (2010) unit root test, a more recently developed and powerful unit root test, which tests the null hypothesis of non-stationarity, i.e. a series has a unit root. Using the NP (2010) test, we find that the null hypothesis of non-stationarity is rejected for aluminium, copper, lead, silver, gas and zinc. This implies that aluminium, copper, lead, silver, gas and zinc are the stationary series and coal, iron, nickel, petroleum and tin are the non-stationary series.

4 Conclusions and implications

In this paper, we examined the time series properties of 11 non-renewable resource price series. The motivation for undertaking this research is that by identifying the actual properties of the data series we are better able to comment on whether or not these resources are subject to increasing degrees of scarcity. If the series is found to be mean reverting, it will indicate that resources are not subject to increasing degrees of scarcity and that is why their price is returned to the mean path and vice versa. To undertake this research, we have employed several tests and grouped them into three heads, namely ‘classical’ unit root tests (which include ADF, PP, DF-GLS and Ng and Perron tests), powerful tests (which include, KPSS, IKPSS, XL and LN tests) and unit root tests which incorporate structural breaks such as the NP (2010) two-break model. To compare the results, we used the KPSS test with two structural breaks developed by Carrion-i-Silvestre and Sansó (2007) and thereby extend the existing literature in the area. We find conflicting results. For example, the results of the Ng and Perron (2001) unit root test show that coal, copper, lead, nickel and zinc are the stationary variables and the remaining seven series are non-stationary. However, results from powerful stationarity tests show that only iron, lead and nickel are strictly stationary and remaining eight series are strictly non-stationary. The results from the Carrion-i-Silvestre and Sansó (2007) stationarity test provide significant evidence to reject the null hypothesis of stationarity for aluminium, iron, tin and zinc, indicating that these variables are not $I(0)$. However, the results obtained through the NP (2010) test show that only aluminium, copper, lead, silver, gas and zinc are stationary variables, i.e. only these variables are $I(0)$. Hence, based on the present study and conclusion drawn from previous results, which used the same data set, we conclude that the results are very much affected by the approach we
use for our analysis. However, we can conclude that for aluminium and zinc it is difficult to reach a conclusion; copper, iron, lead and tin show the stationary path; however, robust evidence is found only for iron and lead.

Further, from the research implication point of view, we found from the plots of the variables (see Figure A1) that most of the series contains drift over time, but one needs to test whether this drift is linear or non-linear such as quadratic. Further, we observe that, greatly, the results obtained and conclusions drawn on the stationary of the series depend not only on the timeframe one uses (as pointed out by Sharma et al., 2009) but also (as we found) on the type of the unit root test or stationarity test one utilises. However, as correctly pointed out by Sharma et al. (2009), the differences in the results greatly depend on the “three characterisations of long run change: (i) episodic mean shifts versus (ii) stationarity around deterministic trends versus (iii) stochastic trends” (p.7). Our robust findings of stationary evidence for iron and lead have strong policy implications for countries which are abundant with these natural resources in designing stabilisation funds for these commodity revenues. Further, one can obtain the reliable forecast for these series even if structural breaks are not incorporated. These variables can also be used for some econometric estimations, but by incorporating the identified breaks. For the other variables which show non-stationary behaviour, identified break dates may also be used in econometric modelling and forecasting.

Acknowledgements

I am grateful to the anonymous referee for constructive comments and suggestions on an earlier draft of the paper without which the paper would have remained flawed and incomplete. I also thank Prof. J.L. Carrion-i-Silvestre, Prof. A. Sansó, Prof. P.K. Narayan, Prof. S. Popp, Prof. L.R. Lima and Prof. B. Neri for sharing the estimation codes of their works.

References

Long-term trends in non-renewable resource


Notes

1 In particular, they found data to be stationary around multiple breaks (i.e., two or more breaks) for almost all the series, except silver.

2 This section is heavily drawn from the study by Narayan and Popp (2010).

3 Here we provide our results for the CC model (which assumes two breaks in level and trend) only in order to compare our results of the KPSS test with the NP test and due to limited space. However, results of the other six specifications of the deterministic model can be obtained from the author upon request.

4 Carrion-i-Silvestre and Sansó (2006a) have recently compared the different procedures to establish a boundary rule showing that the proposal in Sul et al. (2005) is the best in terms of size and power.

5 There are two approaches in the literature of variance stationarity statistics that address the estimation of the break dates. First is the minimum functional to the sequence that results from the computation of the KPSS test for all possible break points as applied by Busetti and Harvey (2007) and Lee and Strazicich (2004). The argument that minimises this sequence is taken as the estimate of the breaking point. Second is the minimisation of the sequence of SSR to estimate the date of the break as suggested by Kurozumi (2002).

6 Results for the Ljung–Box statistic and the ARCH–LM statistic are available upon request. The test statistics related to the powerful tests are also available upon request to the author.
Appendix A

Figure A1 Plots of the variables

Table A1 Seven deterministic specifications of KPSS test

<table>
<thead>
<tr>
<th>Model</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAn</td>
<td>$f(t, T_n, T_{n-1})$</td>
</tr>
<tr>
<td>AA</td>
<td>$\mu + \sum_{i=1}^{2} \theta_i U_{i,t}$</td>
</tr>
<tr>
<td>BB</td>
<td>$\mu + \beta t + \sum_{i=1}^{2} \gamma_i D_{i,t}^*$</td>
</tr>
<tr>
<td>CC</td>
<td>$\mu + \beta t + \sum_{i=1}^{2} \theta_i U_{i,t} + \sum_{i=1}^{2} \gamma_i D_{i,t}^*$</td>
</tr>
<tr>
<td>AB-BA</td>
<td>$\mu + \beta t + \theta_1 U_{1,t} + \gamma_2 D_{2,t}^*$</td>
</tr>
<tr>
<td>AC-CA</td>
<td>$\mu + \beta t + \sum_{i=1}^{2} \theta_i U_{i,t} + \gamma_2 D_{2,t}^*$</td>
</tr>
<tr>
<td>BC-CB</td>
<td>$\mu + \beta t + \theta_2 U_{2,t} + \sum_{i=1}^{2} \gamma_i D_{i,t}^*$</td>
</tr>
</tbody>
</table>
**Table A2** Critical values for ‘classical’ unit root tests

<table>
<thead>
<tr>
<th></th>
<th>ADF</th>
<th>Phillips–Perron</th>
<th>Elliott–Rothenberg–Stock</th>
<th>Ng–Perron</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tau</td>
<td>Z-tau</td>
<td>DF-GLS</td>
<td>P-test</td>
</tr>
<tr>
<td>1pct</td>
<td>-4.04</td>
<td>-4.06364</td>
<td>-3.58</td>
<td>4.26</td>
</tr>
<tr>
<td>Aluminium</td>
<td>-3.45</td>
<td>-3.46023</td>
<td>-3.03</td>
<td>5.64</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.15</td>
<td>-3.15603</td>
<td>-2.74</td>
<td>6.79</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.0367</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Coal</td>
<td>-3.43</td>
<td>-3.4475</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.14858</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.0367</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Copper</td>
<td>-3.43</td>
<td>-3.4475</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.14858</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
<tr>
<td>1pct</td>
<td>-4.04</td>
<td>-4.09086</td>
<td>-3.58</td>
<td>4.26</td>
</tr>
<tr>
<td>Gas</td>
<td>-3.45</td>
<td>-3.47302</td>
<td>-3.03</td>
<td>5.64</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.15</td>
<td>-3.16346</td>
<td>-2.74</td>
<td>6.79</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.04937</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Iron</td>
<td>-3.43</td>
<td>-3.45349</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.15209</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.0367</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Lead</td>
<td>-3.43</td>
<td>-3.4475</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.14858</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
<tr>
<td>1pct</td>
<td>-4.04</td>
<td>-4.08028</td>
<td>-3.58</td>
<td>4.26</td>
</tr>
<tr>
<td>Nickel</td>
<td>-3.45</td>
<td>-3.46806</td>
<td>-3.03</td>
<td>5.64</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.15</td>
<td>-3.16058</td>
<td>-2.74</td>
<td>6.79</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.03734</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Petroleum</td>
<td>-3.43</td>
<td>-3.4478</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.14876</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.0367</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Silver</td>
<td>-3.43</td>
<td>-3.4475</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.14858</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.04766</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Tin</td>
<td>-3.43</td>
<td>-3.45268</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.15162</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
<tr>
<td>1pct</td>
<td>-3.99</td>
<td>-4.0367</td>
<td>-3.46</td>
<td>4.05</td>
</tr>
<tr>
<td>Zinc</td>
<td>-3.43</td>
<td>-3.4475</td>
<td>-2.93</td>
<td>5.66</td>
</tr>
<tr>
<td>10pct</td>
<td>-3.13</td>
<td>-3.14858</td>
<td>-2.64</td>
<td>6.86</td>
</tr>
</tbody>
</table>