# Statistical Assessment of Hotelling's Rule in Forecasting the Crude Oil Price: Threshold Cointegration Approach

Byeongseon Seo and Hojeong Park

Korea University

**Abstract**

The net price of an exhaustible resource rises over time at the interest rate with certainty according to the r-percent rule of Hotelling (1931). Recent movements of crude oil prices are inconsistent with the r-percent rule in that the crude oil futures price far exceeded the spot price. Uncertainty in demand and reserves may affect the expected change in the oil price, and the speculation and storage behavior may lead to the price spikes. This paper provides a statistical assessment of Hotelling's rule in forecasting the crude oil prices. The threshold vector error correction model is employed to allow for nonlinear dynamic adjustment of the crude oil prices. The statistical evidence of nonlinear dynamic adjustment in the futures and spot price relationship is statistically significant. In terms of the forecast power, the threshold VECM forecasts outperform the forecasts based on the r-percent rule and the linear VECM.

JEL Classification: C51; Q40

Keywords: crude oil; Hotelling's rule; nonlinear adjustment; threshold cointegration

---

*Department of Food and Resource Economics, College of Life Science, Korea University, Seoul, 136-701, Korea. E-mail: seomatteo@korea.ac.kr (Seo); hjeongpark@korea.ac.kr (Park). Phone: 82-2-3290-3032. Fax: 82-2-953-0737.

## 1. Introduction

The r-percent rule of Hotelling (1931) implies that the price of an exhaustible resource rises over time at the interest rate with certainty. However, recent movements of crude oil prices show that the futures price far exceeded the spot price, which does not conform to the r-percent rule. Hotelling's rule is based on the assumption of no uncertainty and perfect competition. Uncertainty in demand and reserves may affect the expected change in the oil price. In addition, the speculation and storage behavior along with the perception of a future supply shortage may lead to the price spikes. This paper provides a statistical assessment of Hotelling's rule in forecasting the crude oil prices.

The behavior of commodity prices reveals the stylized facts of extreme volatility, skewed distribution, and high degree of price autocorrelation as discussed in Deaton and Laroque (1992). In particular, the impossibility of carrying negative inventories introduces nonlinearity of the commodity price series. The nonlinearity of the price movement has been suggested by Samuelson (1971) in proving the optimality of competitive storage. As mentioned in Deaton and Laroque (1996), the stylized facts of commodity prices cannot be explained without allowing for the speculation and storage behavior in the commodity market.

Alike the commodities considered in Deaton and Laroque (1996), the crude oil can be stored and carried forward, and thus the behavior of speculative arbitrage generates the equilibrium condition that the expected change in prices is determined by the carry over costs such as interest and shrinkage costs. If the future change in the oil prices is higher than the carryover costs, a positive carryover takes place since it produces net gain. On the other hand, even if the future change in the oil prices is less than the carryover costs, a negative carryover can be limited by the impossibility of carrying negative inventories. Thus, the price dynamics may reveal asymmetric response to the deviation from the equilibrium condition.

The r-percent rule of Hotelling (1931) is closely related to the intertemporal equilibrium condition of Samuelson (1971). Without the shrinkage costs and under certainty, the intertemporal equilibrium condition is equivalent to the r-percent rule. The nonlinear price behavior has been studied in Salant (1983) and Wright and Williams (1984) to explain the speculative attack and the welfare effect of storage, respectively. In the paper, we use the threshold vector error correction model to incorporate nonlinear dynamic adjustment of the crude oil prices.

Figure 1 shows that the spot oil price increased by 56% for a year from January 2007 to December 2007. In addition, the crude oil market was in contango for the periods January 2007 to March 2007 and June 2007 to December 2007 as Figure 2 shows. The futures price was higher than the spot price, and the discounted futures price was about 15% higher than the spot prices in October 2007. This phenomenon may reflect the opportunity of carrying over along with the perception of a future supply shortage. At the same time, the contango in the crude oil market generated an arbitrage opportunity since the combination of buying the spot oil and writing the futures contract could produce riskless net gain. It is natural to think that the arbitrage opportunity is likely to continue to the limit of the global oil storage capacity. As a consequence, the crude oil prices continued to rise in the year 2007 and reached the price spike in July 2008.

On the other hand, backwardation was prevalent during the 1990's. The crude oil futures price was below the spot price for the period. Litzenberger and Rabinowitz (1995) show that backwardation was associated with uncertainty regarding future oil prices. Considine and Larson (2001) suggest that the convenience yield and risk premium were important elements of crude oil price backwardations during the period 1987 to 1997.

Futures price is the price at which both the buyers and sellers are in full agreement to trade oils upon delivery. Therefore, futures price has been acknowledged for the best predictor of future oil prices. However, due to the risk premium, convenience yield, or inventory effects among others, oil futures prices do not necessarily reflect the best forecast measures of future spots (Pindyck, 2001; Considine and Larson, 2001). In the commodity futures market, commodity suppliers write futures contract at a price below the expected future spot price to avoid the price risk according to the hedging theory of Keynes (1930). Previous studies such as Beck (1993) have found that the price risk plays a determinant role in the risk premium, defined as the difference between the futures price and the expected future spot price.

The recent oil price movement is associated with the rocketed futures price of the crude oil market. The empirical observation of the oil market commodity futures price reveals wide and long-lasting violations from the equilibrium condition, which is based on the r-percent rule and also the theory of storage. The price differential between the futures and spot prices increased continuously and did not vanish for long periods. The prolonged deviations from the equilibrium condition can be attributed to the expected risk premium. Thus, we allow for the regime-dependent mean-reverting behavior of the risk premium.

The role of speculation in the commodity futures market can be evaluated by the forecast power since the accuracy of future price forecasts reduces the price adjustment necessary. Fama and French (1987) have evaluated the forecast power in a regression of the future change in spot prices on the basis. The model could not account for the effect of time-varying risk premium. Our model posits the regime-dependent effect of the risk premium in explaining the future change in oil prices.

This paper aims to explore the nonlinear price adjustment and provide a statistical assessment of forecast power contained in the crude oil futures and spot prices relationship. The threshold vector error correction model is employed to assess nonlinear dynamic adjustment. The statistical evidence of nonlinear dynamic adjustment in the futures and spot price relationship is statistically significant. The threshold VECM forecasts outperform the forecasts based on the r-percent rule and the linear VECM.

The surge in the oil and energy prices has attracted a considerable amount of attention in the literature of energy economics. There have been a number of studies examining the relationship between crude oil spot and futures prices, including Abosendra and Baghestani (2004), Bekiros and Diks (2008), Kaufmann et al. (2008), Schwarz and Szakmary (1994). In this paper, we provide empirical evidence of nonlinear price adjustment in the crude oil market, and thereby contribute to the literature.

The paper is organized as follows. The next section provides a stochastic model of the crude oil prices, which explores the price behavior under uncertainty in demand and reserves. Section 3 deals with the econometric methodology. Main empirical results are provided in Section 4.

## 2. Oil Price Dynamics: r-percent Rule

According to Hotelling (1931), the net price of an exhaustible resource will rise over time at the rate of interest. More formally, this principle states $\frac{p_t - p_{t-1}}{p_{t-1}} = r$ or equivalently

$$p_t = p_0 e^{rt},$$

where $p_t$ denotes the net price of exhaustible resources such as the crude oil and $r$ is the interest rate. This implies that, if discounted future price is higher than the spot price, producers would rationally delay current production and expand future production. However, expansion for future production is likely to be costly due to limitations of residual production capacity. For instance, investment decision tends to be delayed when uncertainty prevails in the future (Dixit and Pindyck, 1994). Uncertainty in demand may be a reason for a failure of the r-percent rule by creating slow adjustment for production capacity in oil business. Consequently, uncertainty implies non-equivalence between the net price change and the interest rate. In addition, Considine and Larson (2001) examine how inventory change affects the oil price dynamics.

In the paper, we provide a stochastic model of oil price dynamics and generalize the r percent rule by incorporating uncertainty in demand and reserves as well as the inventory change. Our model allows for the dynamics of inventory and impossibility of carrying negative inventories, and thus extends Pindyck (1980) which consider only demand and reserves uncertainty. The r-percent rule is a special case of our model as discussed below.

We consider the profit maximization problem with stock-dependent extraction cost given the discount rate $r$. For each instant of time, the crude oil is extracted at the rate $q(t)$ and among them the sales is made at the rate $y(t)$. Thus, the oil inventory $I(t)$ changes at the rate $q(t) - y(t)$. The cost function depends on $q(t)$, $y(t)$, and oil reserves under the ground $R(t)$. Thus, the producer's objective is given by

$$Max_{\{q(t),y(t)\}_t^T} E_t[\int_t^T (p(t)y(t) - C(q(t), y(t), R(t), I(t)))e^{-rt} dt] \quad (1)$$

$$dp(t) = \alpha p(t)dt + \sigma_1 p(t)dz_1 \quad (2)$$

$$dR(t) = -q(t)dt + \sigma_2 dz_2 \quad (3)$$

$$dI(t) = (q(t) - y(t))dt \text{ and } I(t) \geq 0 \text{ for all } t. \tag{4}$$

Equation (1) represents the maximization of the present value of the future stream of profits from extracting $\{q(t)\}_t^T$ and sales $\{y(t)\}_t^T$, where $T$ is the terminal point of exhaustion. A diffusion process of the geometric Brownian motion is assumed for the price to incorporate uncertainty in the oil price demand in equation (2), where $z_1$ is the standard Wiener process. The drift and volatility parameters are assumed to be constant. Equation (3) allows for reserves uncertainty with $z_2$ that follows a Wiener process. It is assumed that $z_2$ is independent of $z_1$.[1] The inventory changes as equation (4), and it satisfies impossibility of carrying negative inventories in the same way as in Samuelson (1971). In what follows, time $t$ is suppressed for notational simplicity.

Denote $V(p, R, I)$ as the value function associated with the profit maximization problem presented above. Then, using the Ito's lemma, we have the Hamilton-Jacobi-Bellman equation

$$0 = \max_{q,y}\left[ py - C(q, y, R, I) + V_t + (q - y)V_I + \alpha_1 pV_p + \frac{1}{2}\sigma_1{}^2 p^2 V_{pp} - qV_R + \frac{1}{2}\sigma_2{}^2 V_{RR} \right] \tag{5}$$

The first order conditions with respect to $q$ and $y$ are given as follows:

$$\frac{\partial \pi}{\partial q} = V_R - V_I \tag{6}$$

$$\frac{\partial \pi}{\partial y} = V_I \tag{7}$$

where $\pi = (py - C(q, y, R, I))e^{-rt}$.

We differentiate equation (5) with respect to $R$ and $I$ to get $\partial\pi / \partial R + 1/dt(E_t dV_R) = 0$ and $\partial\pi / \partial I + 1/dt(E_t dV_I) = 0$. Also, using Ito's lemma, we obtain $1/dt(E_t d(\partial\pi / \partial q)) = 1/dt(E_t d(V_R - V_I))$ and $1/dt(E_t d(\partial\pi / \partial y)) = 1/dt(E_t dV_I)$. Thus, we get

$$\frac{1}{dt}E_t d\left(\frac{\partial \pi}{\partial q}\right) + \frac{1}{dt}E_t d\left(\frac{\partial \pi}{\partial y}\right) = -\frac{\partial \pi}{\partial R}. \tag{8}$$

---

[1] We may assume that these two Wiener processes are correlated. Then, there appears an interaction term in equation (5).

First, by applying $\partial\pi/\partial y = \left(p - C_y\right)e^{-rt}$ to (7) and using Ito's lemma along with $E_t\left(dR\right) = -qdt$ and $E_t(dR^2) = \sigma_2 dt$, we have $(1/dt)E_t dC_y = -qC_{yR} + (1/2)\sigma_2{}^2 C_{yRR}$. This result is substituted into $1/dt\left(E_t d\left(\partial\pi/\partial y\right)\right) = 1/dt\left(E_t dV_I\right)$ to obtain

$$\frac{1}{dt}E_t dp = r(p - C_y) + C_I - qC_{yR} + \frac{1}{2}\sigma_2{}^2 C_{yRR} \tag{9}$$

Equation (9) is an r-percent rule directing the optimal price path with respect to sales. Note that $C_I$ is interpreted as the Kaldor convenience yield as in Considine and Larson (2001) by measuring the benefit of holding inventories. Similar steps for the first order condition (6) using $\partial\pi/\partial q = -C_q e^{-rt}$ and $(1/dt)E_t dC_q = -qC_{qR} + (1/2)\sigma_2{}^2 C_{qRR}$ provides

$$C_R - C_I - rC_q - qC_{qR} + \frac{1}{2}\sigma_2{}^2 C_{qRR} = 0. \tag{10}$$

Finally, we have (11) by putting together (9) and (10) into the form of (8):

$$\frac{1}{dt}E_t dp = r\left(p - C_y - C_q\right) + C_R - q\left(C_{qR} + C_{yR}\right) + \frac{1}{2}\sigma_2{}^2\left(C_{qRR} + C_{yRR}\right) \tag{11}$$

Note that, unlike Considine and Larson (2001), the Kaldor convenience yield disappears in (11) since the producer now optimally chooses the extraction rate and sales rate simultaneously so that the benefit of holding inventory is limited. We assume that the marginal cost of sales is independent of the amount of reserves, which implies $C_{yR} = C_{yRR} = 0$. Then, equation (11) can be simplified as follows:

$$\frac{1}{dt}E_t dp = r\left(p - C_y - C_q\right) + C_R - qC_{qR} + \frac{1}{2}\sigma_2{}^2 C_{qRR} \tag{12}$$

Note that equation (12) corresponds to the stochastic version of the r-percent rule. If the cost function does not involve reserves ($C_R = 0$), and the marginal cost is constant such that $C_y = C_q = 0$, then equation (10) can be reduced to the r-percent rule of Hotelling. That is, the oil price evolves in accordance with $1/dtE_t dp = rp$. However, introduction of more generalized cost function brings in the deviation of oil price dynamics from the r-percent rule. We could see that equation (12) relates to the discussion of futures and optimal price path is closely dependent on the property of cost function. For example, under the reasonable assumption of $C_{qRR} > 0$, the change in price tends to increase if uncertainty in reserves, $\sigma_2$, is sufficiently large. Therefore, under the circumstance of high uncertainty on reserve, it is more likely to have contango. Conversely, with less uncertainty, we may have backwardation where the price reveals relatively lower growth rate. It must be noted that, due to the impossibility of negative inventory carryover, backwardation tends to be limited relatively to contango. However, we cannot rule out that the opposite case may hold when we have $C_{qRR} < 0$ although this is less likely. In sum, the change in crude oil prices can be affected by multiple factors such as uncertainty in reserves and inventory dynamics as well as the marginal cost of extraction and sales. As will be analyzed below, thus, the adjustment of oil prices shows the regime-dependent behavior.

## 3. Econometric Methods

This section deals with the econometric models that can be used to estimate the nonlinear dynamic adjustment behavior of the crude oil futures and spot prices. We denote $f_t$ as the oil futures price and $s_t$ as the spot price in logarithms. Also, $g_t$ denotes the discounted futures price. The r-percent rule states that the expected change in the crude oil price is determined by the interest rate, that is,

$$E_t(s_{t+m}) = s_t + r_t m, \tag{13}$$

where $m$ denotes the time to maturity.

Although Hotelling's r-percent rule is deterministic, we allowed for stochastic movement of the price. As in Beck (1993), we define the risk premium $w_t$ as the difference between the futures price $f_t$ and the expected spot price $E_t(s_{t+m})$. Thus, from the r-percent rule we have the futures and spot price relationship as follows:

$$f_t = s_t + r_t m + w_t, \tag{14}$$

where $w_t = f_t - E_t(s_{t+m})$.

The futures and spot price relationship is similar to the intertemporal equilibrium condition of Samuelson (1971) if the shrinkage from carryover is zero. Also, the theory of storage implies the condition of no arbitrage, which is the same as the futures and spot price relationship without the cost of storage and the convenience yield. If we assume rational expectations, the expected spot price is the same as the futures price since the risk premium will be zero. We assume that the risk premium is stationary in the sense that the dependence of the risk premium vanishes eventually.

To assess the price adjustment behavior, we use the vector error correction model (VECM) as follows:

$$\Delta x_t = \mu + \alpha w_{t-1}(\beta) + \sum_{i=1}^{k} \Gamma_i \Delta x_{t-i} + u_t , \qquad (15)$$

where $x_t = (g_t, s_t)'$, $g_t = f_t - r_t m$ and $E_{t-1}(u_t) = 0$.

The long-run relationship is defined as $w_t(\beta) = g_t - \beta s_t$, which should be stationary according to Engle and Granger (1987). In empirical results, the stationarity of the relationship will be analyzed. To estimate the response of the spot prices to the equilibrium condition, we set the cointegration coefficient $\beta$ at unity in the paper. We may allow for unknown cointegrating coefficient, and the results are similar to those with fixed coefficient. If $\beta$ =1, the long-run relationship is the same as the disequilibrium error, which is defined as

$$w_t = g_t - s_t = f_t - s_t - r_t m .$$

The mean-reverting behavior depends on the adjustment vector $\alpha$ in the linear VECM. If it is close to 0, the futures and spot prices do not respond to the disequilibrium error, which renders the disequilibrium error persistent. The linear VECM assumes that the adjustment vector is constant and the response of the spot price to the past deviation is linear.

As Figure 3 shows, the response of the spot price changes becomes strong in the lower tail compared to the response in the middle and in the upper tail. In the lower tail, the spot price is higher than the predicted value of the no arbitrage condition. The spot price responds to the disequilibrium errors. In the mid regime, the arbitrage opportunity cannot be exploited as the arbitrage does not produce net gain because of the storage cost. The upper tail corresponds to the arbitrage opportunity of carrying forward as the futures price exceeds the spot price by the cost of carrying over. The storage is profitable in the upper regime, which generates slow adjustment.

The linear VECM assumes linearity and does not allow for the nonlinear dynamic adjustment behavior. The adjustment behavior depicted in Figure 3 can be explained by a threshold model as follows:

$$
\Delta w_t = \begin{cases} a_1 w_{t-1} + \nu_1 + \sum_{i=1}^{k} c_{1i} \Delta w_{t-i} + v_t \\ a_2 w_{t-1} + \nu_2 + \sum_{i=1}^{k} c_{2i} \Delta w_{t-i} + v_t \\ a_3 w_{t-1} + \nu_3 + \sum_{i=1}^{k} c_{3i} \Delta w_{t-i} + v_t \end{cases} \quad \text{for} \quad \begin{array}{c} w_{t-1} \le \gamma_1 \\ \gamma_1 < w_{t-1} \le \gamma_2 \\ w_{t-1} > \gamma_2 \end{array} \tag{16}
$$

The adjustment coefficients vary depending on the state determined by the arbitrage opportunity and the threshold parameters. In our analysis, a multivariate threshold model is employed to assess the nonlinear dynamic behavior. The threshold vector error correction model (TVECM) has been developed by Hansen and Seo (2002) as a means to combining the long-run relationship and the nonlinear adjustment. Since the empirical regularity shows that the oil spot prices respond to the disequilibrium error asymmetrically, the threshold model may improve the forecasting power.

The threshold vector error correction model (VECM) can be defined as follows:

$$
\begin{aligned} \Delta x_t = & [\mu_1 + \alpha_1 w_{t-1} + \sum_{i=1}^{k} \Gamma_{1i} \Delta x_{t-i}] \times 1(w_{t-1} \le \gamma_1) + \\ & [\mu_2 + \alpha_2 w_{t-1} + \sum_{i=1}^{k} \Gamma_{2i} \Delta x_{t-i}] \times 1(\gamma_1 < w_{t-1} \le \gamma_2) + \\ & [\mu_3 + \alpha_3 w_{t-1} + \sum_{i=1}^{k} \Gamma_{3i} \Delta x_{t-i}] \times 1(w_{t-1} > \gamma_2) + u_t \end{aligned} \tag{17}
$$

where $1(\cdot)$ is the indicator function.

The disequilibrium error $w_{t-1} = g_t - s_t$ and the threshold parameter $\gamma = (\gamma_1, \gamma_2)$ determine three regimes. Regime 1 corresponds to the period when the disequilibrium error is smaller than the threshold parameter $\gamma_1$. Regime 3 is defined as the period when the disequilibrium error is larger than the threshold parameter $\gamma_2$. Regime 2 is the period when the disequilibrium error is in between $\gamma_1$ and $\gamma_2$.

The adjustment vectors are $\alpha_1$, $\alpha_2$, and $\alpha_3$ for each regime and they are regime-dependent. The intercepts and the coefficients of short-run dynamic parameters are also regime-varying.

In our model, the regime is determined by the long-run equilibrium relationship and the threshold parameters. Thus, our model is different from the conventional regime switching model. As the adjustment vectors are allowed to change, our model can explain the asymmetric mean-reverting behavior of the crude oil prices.

The tests for nonlinear dynamic adjustment in the crude oil prices can be based on the following hypotheses:

$H_0 : \delta = 0$ against $H_1 : \delta \neq 0$,

where $\delta = vec(\delta_1, \delta_3)$,

$\delta_1 = vec(\mu_1 - \mu_2, \alpha_1 - \alpha_2, \Gamma_{11} - \Gamma_{21}, \Gamma_{12} - \Gamma_{22}, \ldots, \Gamma_{1k} - \Gamma_{2k})$

$\delta_3 = vec(\mu_3 - \mu_2, \alpha_3 - \alpha_2, \Gamma_{31} - \Gamma_{21}, \Gamma_{32} - \Gamma_{22}, \ldots, \Gamma_{3k} - \Gamma_{2k})$ and $vec(\cdot)$ is the operator of vectorization,

Under the null hypothesis, the TVECM reduces to the linear VECM, and then the disequilibrium error follows a linear mean-reverting behavior. If we fix the threshold parameters, the TVECM can be estimated by linear regression and the tests for nonlinear mean reversion can be based on the LM statistic as follows:

$$LM(\gamma) = \hat{\delta}'[Var(\hat{\delta})]^{-1}\hat{\delta},$$

where $Var(\hat{\delta})$ is the estimated variance of $\hat{\delta}$, and the heteroskedasticity-robust covariance estimator is used since the data contains conditional heteroskedasticity.

The threshold parameter $\gamma$ cannot be identified under the null hypothesis, and as a result the standard testing methods cannot be applied. We use the Sup-LM statistic which does not depend on the nuisance parameter.

$$SupLM = \sup_{\gamma \in [\gamma_L, \gamma_U]^2} LM(\gamma)$$

The lower limit of the threshold parameter $\gamma_L$ is the $p$ -th percentile of the disequilibrium error, and $\gamma_U$ is the $(1-p)$ -th percentile. The truncation parameter $p$ is set at 0.10, and the main results do not change a lot at other values such as 0.05.

The asymptotic distribution of the SupLM statistic follows a nonstandard distribution as shown by Hansen and Seo (2002). Hansen and Seo suggest statistical inference using the bootstrapping p-values. The null hypothesis can be rejected if the bootstrapping p-values are smaller than the size chosen.

## 4. Main Results

In the empirical work, we use the daily data on the crude oil spot and futures prices and the 3-month US Libor rate. The crude oil prices are for West Texas Intermediate, the U.S. benchmark grade. The spot prices of crude oil are closing prices on the New York Mercantile Exchange (NYMEX). The future prices of crude oil are NYMEX settlement prices for the nearby one-month ahead contracts, which show the largest and dominant trading volume. The data set is obtained from the Petronet database which is provided by Korea National Oil Corporation (KNOC) for the sample period January 2003 through December 2011.

We first analyze the basis defined as $f_t - s_t$, where $f_t$ is the futures price and $s_t$ is the spot price in logarithms. Also, the discounted futures prices $g_t = f_t - r_t m$ and the discounted basis $w_t = g_t - s_t$ are defined to examine the time series characteristics. As Table 1 shows, the distribution of the basis is skewed to the right, which reflects the contango for the period January 2007-December 2007. The leptokurtic behavior, serial correlation, and conditional heteroskedasticity exist in the basis and the discounted basis. As discussed in Deaton and Laroque (1992), the commodity prices display the stylized facts such as skewed distribution, violent explosions, and high degree of autocorrelation. Most facts found in Table 1 are similar to those of Deaton and Laroque.

[Table 1] Descriptive Statistics of the Basis and the Disequilibrium Error

| | Mean | S.D. | Skewness | Kurtosis | Q-Statistic | ARCH-LM |
|---|---|---|---|---|---|---|
| $f_t - s_t$ | 0.003614 | 0.024446 | 2.077788 | 18.08738 | 4967.7<br>[0.000] | 712.0817<br>[0.000] |
| $g_t - s_t$ | 0.001115 | 0.024264 | 2.000328 | 18.33265 | 4825.8<br>[0.000] | 662.9644<br>[0.000] |

* Q-statistics and ARCH-LM statistics are computed at the AR lag length 6. The p-values are in the brackets.

Table 2 shows the ADF unit root tests. The unit root hypothesis of futures, spot, and discounted futures prices cannot be rejected by the ADF unit root test. The optimal AR lag lengths picked by the BIC are all 2.

[Table 2] ADF Unit Root Tests

| | AR Lag | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $f_t$ | Statistic | -3.022788* | -2.843340 | -2.861757 | -2.884073 | -2.742394 |
| | p-value | 0.1260 | 0.1817 | 0.1753 | 0.1678 | 0.2194 |
| $s_t$ | Statistic | -2.989510* | -2.952697 | -2.977193 | -3.029716 | -2.883903 |
| | p-value | 0.1352 | 0.1460 | 0.1388 | 0.1241 | 0.1678 |
| $g_t$ | Statistic | -3.032979* | -2.852883 | -2.871065 | -2.893361 | -2.751263 |
| | p-value | 0.1233 | 0.1784 | 0.1722 | 0.1647 | 0.2159 |

*: BIC-minimizing AR lag length

As Table 3 shows, the likelihood ratio statistics reject the null hypothesis of no cointegration at the 5% size. The null hypothesis of one cointegration rank maintains for each VAR lag lengths. Thus, we find a long-run relationship between the futures and the spot prices by using Johansen's (1988) cointegration tests. For example, at the VAR lag-length 6 picked by BIC, the likelihood ratio statistic for a bivariate cointegration between the discounted futures and the spot prices is 47.16 which exceeds the asymptotic critical value at the 5% size. The null hypothesis of one cointegrating relationship cannot be rejected for any combination of the futures and spot prices for all VAR lag lengths.

[Table 3] Johansen's Cointegration Tests

| | VAR Lag | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $(f_t, s_t)$ | LR(rank=0) | 173.2332 | 131.3756 | 98.16077 | 76.19394 | 59.49699 | 44.37949 |
| | p-value | 0.0001 | 0.0001 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| | LR(rank=1) | 3.166684 | 3.160999 | 3.201762 | 3.278988 | 2.980172 | 2.879124 |
| | p-value | 0.0752 | 0.0754 | 0.0736 | 0.0702 | 0.0843 | 0.0897 |
| $(g_t, s_t)$ | LR(rank=0) | 182.2092 | 138.7381 | 103.8872 | 80.74116 | 63.19087 | 47.16400 |
| | p-value | 0.0001 | 0.0001 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| | LR(rank=1) | 3.162326 | 3.154778 | 3.193899 | 3.270234 | 2.969388 | 2.868173 |
| | p-value | 0.0754 | 0.0757 | 0.0739 | 0.0705 | 0.0849 | 0.0903 |

Table 4 shows the estimates of the linear vector error correction model for the discounted futures and spot oil prices. The VAR lag length chosen is 6 by BIC. Standard errors are based on the heteroskedasticity-consistent covariance estimator.

[Table 4] Estimates of Linear VECM

| | $\Delta g_t$ | | $\Delta s_t$ | |
|---|---|---|---|---|
| | Coefficients | standard errors | coefficients | standard errors |
| January 2003-December 2011 | | | | |
| $\hat{\alpha}$ | 0.006925 | 0.02888 | 0.113429 | 0.02721 |
| $\hat{\mu}$ | 0.000600 | 0.00059 | 0.00005 | 0.00055 |
| Log likelihood = 10970.17 | | | | |
| January 2007-July 2008 | | | | |
| $\hat{\alpha}$ | -0.007187 | 0.03220 | 0.045690 | 0.02822 |
| $\hat{\mu}$ | 0.002275 | 0.00149 | 0.000257 | 0.00130 |
| Log likelihood = 1938.895 | | | | |

The adjustment coefficient of the spot price is positive and statistically significant for the sample period January 2003-December 2011. The adjustment coefficient of the discounted futures price is not statistically significant. The persistence of a shock in the disequilibrium error can be measured by its half life. The half life of a shock is estimated at 4.40. For the same period January 2007-July 2008, the adjustment coefficient of the spot price is significant although its size becomes smaller. During the period January 2007-July 2008, the crude oil price continued rising, which resulted sluggish adjustment to the r-percent rule.

[Table 5] Testing for Nonlinear Adjustment

| VAR Lag | SupLM Statistic | p-value | 5% critical value |
|---|---|---|---|
| 1 | 46.332 | 0.000 | 35.312 |
| 2 | 53.204 | 0.004 | 45.186 |
| 3 | 74.311 | 0.000 | 56.275 |
| 4 | 100.307 | 0.000 | 65.435 |
| 5 | 102.589 | 0.000 | 75.515 |
| 6 | 111.322 | 0.000 | 85.783 |

Next, we use the threshold vector error correction model proposed by Hansen and Seo (2002). The tests for nonlinear dynamic adjustment support the hypothesis of nonlinear adjustment in the futures and spot prices relationship as Table 5. For example, the SupLM statistic for the tests of nonlinear mean reversion is calculated at 85.783 with a bootstrapping p-value of 0.000. The tests are based on the threshold VECM with a VAR lag length of 6, which is picked by BIC, and the trimming parameter p=0.10. The bootstrapping p-values are calculated on the linear error correction model with 1,000 bootstrapping replications. The SupLM statistics indicate in favor of nonlinear price dynamics for other values of the VAR lag length even if we change the lag order.

This result implies that the change in oil prices respond to the disequilibrium error differently depending on the states, which are determined by the futures and spot price equilibrium condition. Although disequilibrium error in the crude oil market exists, it may come from uncertainty in demand and reserves. Thus, the forecast power of the r-percent rule can be affected by nonlinear price dynamics.

Table 6 shows the estimation result of the threshold VECM, which is estimated by maximum likelihood estimation at the VAR lag-length 6. Standard errors are calculated from the heteroskedasticity-robust covariance estimator. The trimming parameter p is set at 0.10. The threshold estimates are -0.00636, 0.01846, and P($w_{t-1} \le \hat{\gamma}_1$), P($\hat{\gamma}_1 < w_{t-1} \le \hat{\gamma}_2$), and P($w_{t-1} > \hat{\gamma}_2$) are estimated at 0.1003, 0.7873, and 0.1125, respectively. The adjustment coefficient of the spot prices is positive and significant in regime 1. In regime 2, the adjustment coefficient of the spot prices is positive, but its magnitude is not significant.

[Table 6] Estimates of Three-Regime Threshold VECM

| | Regime 1 $w_{t-1} \le \gamma_1$ | | Regime 2 $\gamma_1 < w_{t-1} \le \gamma_2$ | | Regime 3 $w_{t-1} > \gamma_2$ | |
|---|---|---|---|---|---|---|
| | $\Delta g_t$ | $\Delta s_t$ | $\Delta g_t$ | $\Delta s_t$ | $\Delta g_t$ | $\Delta s_t$ |
| January 2003-December 2011 | | | | | | |
| $\hat{\alpha}$ | -0.01448 | 0.41669 | -0.11834 | 0.31467 | 0.03401 | 0.16445 |
| s.e.( $\hat{\alpha}$ ) | 0.19335 | 0.19464 | 0.19444 | 0.19542 | 0.07112 | 0.06575 |
| $\hat{\mu}$ | 0.00529 | 0.00114 | 0.00001 | 0.00054 | 0.00054 | -0.00719 |
| s.e.( $\hat{\mu}$ ) | 0.00384 | 0.00312 | 0.00065 | 0.00062 | 0.00437 | 0.00401 |
| $\hat{\gamma}_1$ = -0.00636, $\hat{\gamma}_2$ = 0.01846<br>P( $w_{t-1} \le \hat{\gamma}_1$ ) = 0.10027, P( $\hat{\gamma}_1 < w_{t-1} \le \hat{\gamma}_2$ ) = 0.78726, P( $w_{t-1} > \hat{\gamma}_2$ ) = 0.11247<br>Log likelihood = 15201.33 | | | | | | |
| January 2007-July 2008 | | | | | | |
| $\hat{\alpha}$ | -0.56360 | 0.06943 | 0.05249 | 0.08100 | -0.54096 | 0.37247 |
| s.e.( $\hat{\alpha}$ ) | 0.22309 | 0.15706 | 0.04776 | 0.04057 | 0.37149 | 0.22497 |
| $\hat{\mu}$ | -0.00156 | -0.00079 | 0.00039 | -0.00103 | 0.04158 | -0.04033 |
| s.e.( $\hat{\mu}$ ) | 0.00300 | 0.00333 | 0.00228 | 0.00201 | 0.03264 | 0.02083 |
| $\hat{\gamma}_1$ = -0.00249, $\hat{\gamma}_2$ = 0.08954<br>P( $w_{t-1} \le \hat{\gamma}_1$ ) = 0.28042, P( $\hat{\gamma}_1 < w_{t-1} \le \hat{\gamma}_2$ ) = 0.61640, P( $w_{t-1} > \hat{\gamma}_2$ ) = 0.10317<br>Log likelihood = 2661.66 | | | | | | |

For the sample period January 2003 to December 2011, the adjustment coefficients of the discounted futures price are negative, but not significant in Regimes 1 and 2. The discounted futures price does not respond strongly to the disequilibrium errors in each regime. Figure 3 shows the response function of the spot and futures prices to the disequilibrium error. The response function is based on the estimates of the intercept and the adjustment vector in each regime given the other short-run dynamics. Figure 4 shows the estimated responses of oil prices to the disequilibrium error. In Regime 1, the spot price increases as the lagged discounted basis increases, and thus the basis provides information on the future change in the crude oil price. In Regime 2, the futures and spot prices do not responds to the disequilibrium error strongly. The disequilibrium errors tend to persist. In Regime 3, the adjustment coefficient of the spot price is not as large as in Regime 1, and the futures price responds positively to the disequilibrium error. Thus, the mean-reverting process becomes active in Regime 1, and the adjustment process in Regime 3 is likely to be weak. Thus, the oil prices reveal the regime-dependent adjustment behavior, and the information content of the equilibrium condition appears regime-specific.

Threshold estimates suggest that the crude oil prices adjust rapidly to the long-run equilibrium, determined by oil prices, in an asymmetric manner, when disequilibria are negative. The dynamic adjustment of spot prices becomes faster when the futures-spot price spread widens and the discounted basis is below a critical threshold. On the other hand, the adjustment of spot prices becomes slow when the discounted basis is above the upper threshold. This result indicates that the contango, if exists, tends to continue longer than the backwardation. The perception of future supply shortage generates the arbitrage opportunity of buying the spot and selling the futures and the opportunity is likely to continue to the limit of the global oil storage capacity. On the other hand, the backwardation does not last long since it is impossible to carry negative inventories.

Table 6 also shows the estimation results for the period January 2007 to July 2008, when the oil prices soared and the contango prevailed. The futures price responds to the disequilibrium error in Regime 1 and Regime 3 while the spot price does not respond to the disequilibrium errors.

Table 7 compares the predictive accuracy of the forecasts based on the r-percent rule, linear VECM, and threshold VECM to the baseline forecasts based on random walk. The forecasts based on the r-percent rule do not show improvement in forecasting the crude oil prices compared to the random walk forecasts.

[Table 7] Predictive Accuracy

| | Random Walk | r-percent Rule | Linear VECM | Threshold VECM |
|---|---|---|---|---|
| | January 2003-December 2011 | | | |
| RMSE | 0.026958<br>(1.0000) | 0.02693<br>(0.9990) | 0.025486<br>(0.9818) | 0.024941<br>(0.9426) |
| MAE | 0.018862<br>(1.0000) | 0.01882<br>(0.9978) | 0.018299<br>(0.9702) | 0.018127<br>(0.9610) |
| | January 2007-July 2008 | | | |
| RMSE | 0.02149<br>(1.00000) | 0.02099<br>(0.9767) | 0.01966<br>(0.9148) | 0.01844<br>(0.8581) |
| MAE | 0.01538<br>(1.00000) | 0.01504<br>(0.9779) | 0.01359<br>(0.8836) | 0.01298<br>(0.8440) |

*The numbers in the parentheses denote relative accuracy compared to the random walk forecasts.

The root mean squared error (RMSE) of the threshold VECM forecasts is about 6% lower than that of the random walk forecasts while the linear VECM forecasts reduce about 2% of the RMSE for the sample period January 2003 to December 2011. In the same model, the mean absolute error (MAE) of the threshold VECM forecasts is about 4% lower than that of the baseline forecasts while that of the linear VECM is 3% lower than that of the forecasts by the random walk hypothesis. The root mean squared error (RMSE) and the mean absolute error (MAE) are defined for one-step ahead forecast errors.

For the sample period January 2007 to July 2008, the predictive accuracy of the threshold VECM forecasts improves by 14% and 16% in terms of RMSE and MAE, respectively, compared to the forecasts based on the random walk hypothesis. The r-percent rule does not improve prediction accuracy greatly compared to the baseline forecasts. For the period, the crude oil prices increased sharply and the discounted futures price far exceeded the spot oil price. Thus, the forecast power improves when we allow for nonlinear price adjustment.

Table 8 shows the tests for equality of forecast accuracy between the forecasts based on the random walk, r-percent rule, linear VECM, and threshold VECM. The statistics are computed by the Morgan-Granger-Newbold tests of which the null hypothesis of equal forecast accuracy is equivalent to zero correlation between the transformed prediction errors. For the sample period January 2003-December 2011, the equality of forecast accuracy between the random walk forecasts and the r-percent rule forecasts cannot be rejected. The equality of prediction accuracy between the forecasts based on the threshold VECM and other forecasts can be rejected. This result shows that the threshold VECM forecasts improve the prediction accuracy compared to the r-percent rule and the linear VECM forecasts.

[Table 8] Morgan-Granger-Newbold Tests for Equality of Forecast Accuracy

| | Random Walk | r-percent Rule | Linear VECM |
|---|---|---|---|
| | January 2003-December 2011 | | |
| r-percent Rule | 1.33935 [0.18069] | | |
| Linear VECM | 5.27387 [0.00000] | 5.26654 [0.00000] | |
| Threshold VECM | 6.74965 [0.00000] | 6.74448 [0.00000] | 3.78847 [0.00016] |
| | January 2007-July 2008 | | |
| r-percent Rule | 2.06281 [0.03989] | | |
| Linear VECM | 3.76609 [0.00020] | 3.73931 [0.00022] | |
| Threshold VECM | 5.97157 [0.00000] | 5.95277 [0.00000] | 4.01812 [0.00007] |

*The p-values are in the brackets.

## 5. Conclusion

This paper assesses the predictive performance of Hotelling's rule on the future change in crude oil prices. We find that the forecast power can be improved by allowing for nonlinear regime-dependent specification, where the regimes are determined by the equilibrium condition. The crude oil spot prices respond strongly to the past disequilibrium error when the disequilibrium error belongs to the lower regime. Our results are consistent with the optimality property of competitive storage proposed by Samuelson (1971), which generates a nonlinear first-order Markov price process. Also, our results support the optimal adjustment model of Pindyck (1980), which considers uncertainty in demand and reserves. The expected rate of change in prices can be affected by the uncertainty in reserves when the extraction cost is nonlinear in reserves.

Also, the main results are suggestive of the threshold vector error correction model, which is consistent with the stylized facts of the crude oil prices. The threshold cointegration model outperforms the linear VECM and the r-percent rule, and thus it can be used in forecasting the crude oil prices.

**References**


Abosendra, S., Baghestani, H., 2004, On the predictive accuracy of crude oil production futures prices, Energy Policy 32 (12), 1389–1393.

Bopp, A.E., Sitzer, S., 1987, Are petroleum futures prices good predictors of cash value? The Journal of Futures Markets 7, 705–719.

Beck, S.E., 1993, A rational expectations model of time varying risk premia in commodity futures markets: theory and evidence, International Economic Review 34, 149-168.

Bekiros, S.D. and C.G.H. Diks, 2008, The relationship between crude oil spot and futures prices: cointegration, linear and nonlinear causality, Energy Economics 30, 2673-2685.

Brown, G.M., Field, B., 1978, The implications of alternative measures of natural resource scarcity, Journal of Political Economy 86, 229-243.

Clark, T.E., McCracken, M.W., 2001. Tests of equal forecast accuracy and encompassing for nested models. Journal of Econometrics 105, 85–110.

Considine, T.J., Larson, D.F., 2001, Uncertainty and the convenience yield in crude oil price backwardation, Energy Economics 23, 533-548.

Crowder, W.J., Hamed, A., 1993. A cointegration test for oil futures market efficiency. Journal of Futures Markets 13, 933–941.

Dasgupta, P.S., Heal, G., 1974, The optimal depletion of exhaustible resources, Review of Economic Studies 41, 3-28.

Deaton, A. and G. Laroque, 1992, On the behaviour of commodity prices, Review of Economic Studies 59, 1-23.

Deaton, A., Laroque, G., 1996, Competitive storage and commodity price dynamics, Journal of Political Economy 104( 5), 896-923.

Dees, S., Karadeloglou, P., Kaufmann, R.K., Sanchez, M., 2007, Modelling the world oil market, assessment of a quarterly econometric model, Energy Policy 35, 178–191.

Dickey, D.A., Fuller, W.A., 1979, Distribution of the estimators for autoregressive time series with a unit root, Journal of American Statistics Association 74, 427–431.

Dixit, A., Pindyck, R., 1994, Investment under Uncertainty, Princeton University Press.

Engle, R.E., Granger, C.W.J., 1987, Cointegration and error-correction: representation, estimation, and testing, Econometrica 55,251–276.

ENI Spa, 2006. World Oil and Gas Review.

Fama, E., French, K., 1987, Commodities futures prices: some evidence on forecast power, premiums and the theory of storage, Journal of Business 60, 55-73.

Gibert, R.J., 1979, Optimal depletion of an uncertain stock, Review of Economic Studies 46, 45-57.

Hansen, B. and B. Seo, 2002, Testing for two-regime threshold cointegration in vector error correction models, Journal of Econometrics 110, 293-318.

Houthakker, H.S., 1957, Can speculators forecast prices?, Review of Economics and Statistics 39, 143-151.

Hotelling, H., 1931, The economics of exhaustible resources, Journal of Political Economy 39, 137-175.

Johansen, S., 1988, Statistical analysis of cointegration vectors, Journal of Economic and Dynamics and Control 12, 231–254.

Kaufmann, R.K., Cleveland, C.J., 2001, Oil production in the lower 48 states: economic, geological, and institutional determinants, Energy Journal 22, 27–49.

Kaufmann, R.K., Dees, S., Karadeloglou, P., Sanchez, M., 2004, Does OPEC matter? An econometric analysis of oil prices, Energy Journal 25 (4), 67–90.

Kaufmann, R.K., S. Dees, A. Gasteuil, M. Mann., 2008, Oil prices: The role of refinery utilization, futures markets and non-linearities, Energy Economics, forthcoming.

Kawai, M., 1983, Price volatility of storable commodities under rational expectations in spot and futures markets, International Economic Review 24, 435-459.

Kilian, L., 2008, A comparison of the effects of exogenous oil supply shocks on output and inflation in the G7 countries, Journal of European Economic Association 6 (1), 78–121.

Morgan, W.A., 1939, A test for significance of the difference between the two variances in a sample from a normal bivariate population, Biometrika 31, 13-19.

Muth, J., 1961, Rational expectations and the theory of price movements, Econometrica 29, 315-335.

Litzenberger, R.H., Rabinowitz, N., 1995, Backwardation in oil futures markets: theory and empirical evidence, Journal of Finance 50(5), 1517-1545.

Moosa, I.A., Al-Loughani, N.E., 1995, The effectiveness of arbitrage and speculation in the crude oil futures market, Journal of Futures Markets 15, 167–186.

Pindyck, R.S., 1978, Gains to producers from the cartelization of exhaustible resources, Review of Economics and Statistics 60, 238-251.

Pindyck, R.S., 1980, Uncertainty and exhaustible resource markets, Journal of Political Economy 88, 1203-1225.

Pindyck, R.S., 2001, The dynamics of commodity spot and futures markets, Energy Journal 22, 1-29.

Regnier, E., 2007, Oil and energy price volatility, Energy Economics 29, 405–427.

Salant, S.W., 1983, The vulnerability of price stabilization schemes to speculative attack, Journal of Political Economy 91, 1-38.

Samuelson, P.A., 1971, Stochastic speculative prices, Proceedings of the National Academy of Science 68, 335-337.

Schwarz, T.V., Szakmary, A.C., 1994. Price discovery in petroleum markets: Arbitrage, cointegration, and the time interval of analysis. The Journal of Futures Markets 14, 147–167.

Serletis, A., Banack, D., 1990, Market efficiency and cointegration: an application to petroleum market, Review of Futures Markets 9, 372–385.

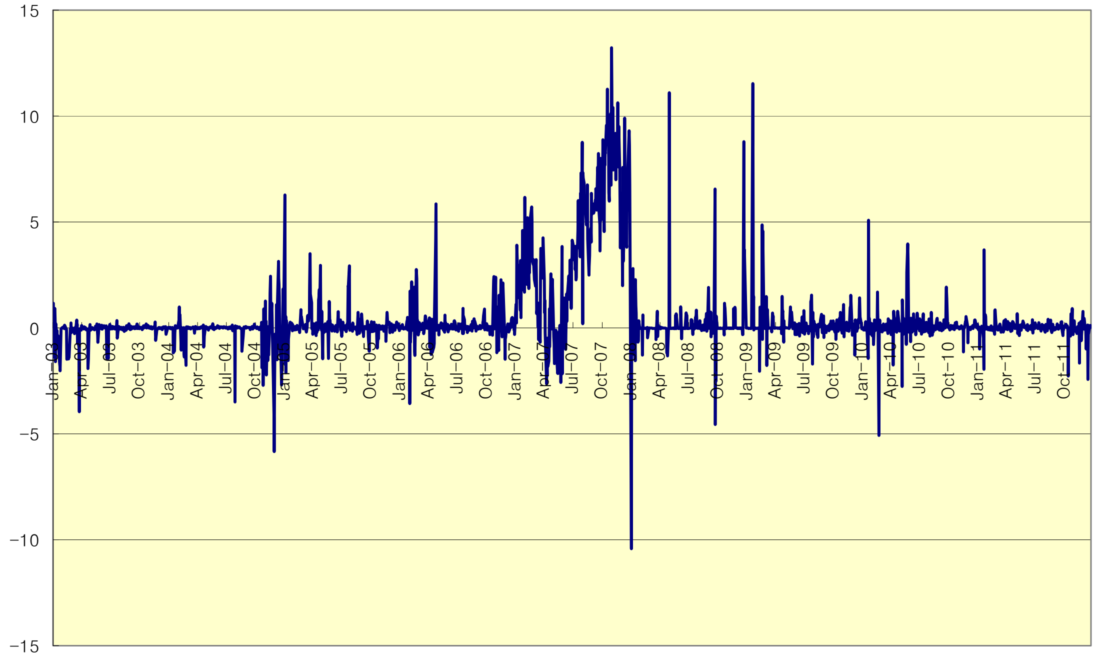
Silvapulle, P., Moosa, I.A., 1999, The relationship between spot and futures prices: evidence from the crude oil market, Journal of Futures Markets 19, 175–193.

Turnovsky, S., 1983, The determination of spot and futures prices with storable commodities, Econometrica 51, 1363-1387.

Wirl, F., Kujundzic, A., 2004, The impact of OPEC conference outcomes on world oil prices 1984–2001, Energy Journal 25 (1), 45–62.

[Figure 1] Crude Oil Spot Price


[Figure 2] Basis



[Figure 3] Response of Spot Price Changes to the Disequilibrium Errors