SNF - WP-- 2/02

Working Paper No. 2/02

Lead Lag Relationships between Futures and Spot Prices

by

Frank Asche
Atle G. Guttormsen

SNF-project No. 7220:
Gassmarkeder, menneskelig kapital og selskapsstrategier

The project is financed by The Research Council of Norway

INSTITUTE FOR RESEARCH IN ECONOMICS AND BUSINESS ADMINISTRATION

BERGEN, JANUARY 2002

ISSN 0803-4028

Lead Lag Relationships between Futures and Spot Prices

by

Frank Asche

Stavanger University College and Centre for Fisheries Economics, Norwegian School of Economics and Business Administration, Box 2557 Ullandhaug, N-4091 Stavanger, Norway. Fax: (47) 51 83 17 50; email: frank.asche@snf.no

and

Atle G. Guttormsen

(Corresponding author)

Department of Economics and Social Sciences, Agricultural University of Norway, Box 5033, N-1432 Aas, Norway. Fax: +47 6494 3012; email: atle.guttormsen@ios.nlh.no

Frank Asche is a professor at Stavanger Universtity College and Centre for Fisheries Economics, Norwegian School of Economics and Business Administration. Atle G. Guttormsen is an associate research professor at the Department of Economics and Social Sciences at the Agricultural University of Norway.

Lead Lag Relationships between Futures and Spot Prices

Abstract

In this paper we examine the relationship between spot and futures prices. This is traditionally done by testing for cointegration with the Engle and Granger methodology, before one specifes an error correction models in order to draw inference about causality. This approach, although appealing for its simplicity, is problematic on at least two accounts. First, the approach is only valid given an exogeneity assumption, which is what one wants to test, and second, given that there are several contracts with different times to expiration, bivariate specifications cannot capture all the relevant information. We show that both problems can be avoided if the tests are carried out in a multivariate framework like the Johansen test. An empirical application is carried out on futures prices for gas oil. Findings indicate that futures prices leads spot prices, and that futures contracts with longer time to expiration leads contracts with shorter time to expiration.

Key Words: Futures, spot, causality, multivariate cointegration, exogeneity.

Introduction

The existence of price discovery, market stability and market efficiency associated with spot and futures markets has been important topics since the genesis of futures markets more than 100 years ago. Numerous papers have examined the relationship between spot and futures prices for various types of commodities as well as for financial assets. Empirical evidence to date is mixed, although a majority of studies indicate that future markets have a price discovering role. Recent papers examining the price discovery role and the lead lag relationship between futures and spot prices have to a large extent followed the two-step procedure outlined in Quan (1992), which is based on the price series being nonstationary.[i] The first step is to test the existence of a long-run relationship between the spot and futures prices by investigating whether the data series are cointegrated. If the first step reveals a long-run relationship, then causality (lead-lag) can be tested to examine the discovery role of futures prices. If no long-run relationship is revealed, the investigation comes to an end because the two times series are generated completely independent (Quan 1992).

When investigating the relationship between spot and future prices, the Engle and Granger test (Engle and Granger, 1987) has been the most common tool to test for cointegration. Conditioned on the existence of a long-run relationship, single equation error correction models (ECM) have been specified in order to draw inference about causality. This approach, although appealing for its simplicity, is problematic on at least two accounts. First, single equation ECMs are only valid given an exogeneity assumption (Banjeree *et al*, 1993). However, this is what one wants to test. Moreover, given that there are several contracts with different times to expiration, bivariate specifications cannot in general capture all the relevant information. This paper show that both problems can be avoided if the tests are carried out in a multivariate framework like the Johansen test (Johansen, 1988, 1991). The first issue can

then be addressed by modeling the relationship between the price of a future contract and the spot price as a bivariate system. The exogeneity assumption underlying any single equation model can then be tested in addition to the existence of a long run relationship. To take into account the fact that at any time several contracts are traded for the same product, a multivariate system must be specified. It is then possible to test several hypothesis with respect to the price discovery process. For instance, in addition to test only whether futures prices lead spot prices, it is of interest to examine whether some of the contracts lead others. Moreover, with the structure most commonly observed for futures prices, it can be shown that at most one price/contract can be leading in the term structure.

Another issue that is often inadequately addressed in the literature is the form of the long-run relationships between the future and spot prices, and particularly if basis is constant so that the prices move proportionally to each other. Often this is assumed, at other times one just regress the relationship without paying attention to this issue, or one investigate the relationships under both conditions. This is because the favored method, Engle and Granger cointegration tests and ECM models, does not allow statistical inference on the parameters in the long-run relationship. This problem can also be avoided in the Johansen framework, as exploited in bivariate systems by Kellard, Newbold, Rayner and Ennew (1999) and Haigh (2000). These tests are also easily extended to a multivariate framework. With the possibility to test restrictions on the parameters in the system, one can test for constant basis as well as for price leadership.

The approach will be used to study futures on Gas oil. Price leadership in the futures market for gas oil and other oil derivatives as well as the crude price has been the focus in a number

of studies including Herbst, McCormack and West (1987), Kawaller, Koch and Koch (1987), Chan (1992), Quan (1992), Schwarz and Szakmary (1994), Moosa and Alloughani, (1994), Gülen (1998), Girma and Paulson (1999) and Silvapulle and Moosa (1999). However there is also conflicting evidence with respect to price leadership. The data used in this paper is from the International Petroleum Exchange (IPE). The gas oil contract was launched as the IPE's first futures contract in 1981, and developed rapidly into a benchmark for spot middle distillate across north-west Europe and beyond. The IPE's Gas Oil futures contract is a highly flexible and liquid contract and is often referred to as heating oil in Europe or the USA. In June 2001, volumes of Gas Oil futures traded reached a daily high of 60,639 lots. IPE claims that the Gas Oil futures have become very closely associated with the physical market.

The paper is organized as follows: The following section will give an exposition of the standard theory on the spot-futures price relationship, in section three multivariate empirical specifications are presented and discussed, whereas the data is presented in section four. Section five present the result and some concluding remarks are presented in the last section

Lead-lag relationships between futures and spot prices.

There are basically two views on the price formation process for commodity futures prices. In the first the intertemporal relationship between cash and futures prices of continuously storable commodities is explained by the cost of carry for the commodity (Kaldor, 1939; Working 1948, 1949; Brennan, 1958; Telser, 1958). In the second view one splits the futures price into an expected risk premium and a forecast of a future spot price (Cootner, 1960; Dusak, 1973; Breeden, 1980; Hazuka, 1984).

The theory of intertemporal relationship between cash and futures prices can be explained briefly with the cost of carry condition as a starting point. This condition can, for continuously storable commodities be formulated as

$$F_t^T \geq S_t e^{(r_t+w)(T-t)} \tag{1}$$

where t is the current date, T is the futures contract expiration date, w a storage cost, r is the riskless money rate of interest at t, F_t^T is the future price at t, and S_t is the spot price at t. This condition says that if you want a commodity at some time T, you can either buy a future contract with delivery at time T, or you can buy the commodity in the spot market and store it until T. A disparity between the left-hand and right-hand side of cost-of-carry equation might give rise to an arbitrage opportunity. Arbitragers can then go long in the commodity and short the futures contract, and hence lock in a secure payoff. From equation (1) it seems like the conditions also should hold the other way around, i.e. that the futures price never should be less than the spot price plus storage and interest cost. However this is more problematic since one can argue that there is a value associated with having the commodity. This value, the convenience yield, is based in the fact that having the commodity in stock provides flexibility regarding for instance unexpected demand.

Supporters of the second view on the price formation process for commodity futures prices argues that the basis can be expressed as a sum of an expected premium and an expected change in the spot price

$$F_t^T - S_t = E_t\left[P(t,T)\right] + E_t\left[S_T - S_t\right] \tag{2}$$

here the expected premium $E_t\left[P(t,T)\right]$ is defined as the bias of the future price as a forecast of the future spot price.

$$E_t\left[P(t,T)\right] = F_t^T - E_t\left[S_T\right] \quad (3)$$

Fama and French (1988) argue that the theory of storage, equation (1) and equation (2) are alternative but not competitive views of the basis, and that variation in expected change in the spot price in (2) translates into variation in the interest rate and the marginal storage cost in (1). Both theories imply that there should be a long run stable relationship between spot and futures prices. In addition, for the future price to be an unbiased predictor of subsequent spot price, i.e., $E_t\left[P(t,T)\right]$ equals zero, the future price should lead the spot price (Garbade and Silber, 1983).

Several other more *ad hoc* arguments for the futures price to lead the spot price can be found in the literature.[ii] Silvapulle and Moosa (1999) argue that futures prices respond to new information more quickly than the latter due to lower transaction costs and ease of shorting, while Newberry (1992) postulate that futures market provide opportunities for market manipulation. According to this argument the futures market can be manipulated either by the larger at the expense of the smaller or by the better informed at the expense of the uninformed. There are also arguments for the opposite hypothesis, that spot price lead futures prices (Silvapulle and Moosa,1999; Quan, 1992; Moosa, 1996). Moosa (1996) presents a model where the change in the spot price will trigger action from the three kinds of market participants, and these actions will subsequently change the futures price.

The theories reviewed above have in common that they are bivariate specifications of the relationship between the spot price and the future price. They do not take into account the fact that for most futures, several contracts are traded at the same time. For a trader in the market, buying a futures contract with expiration at time t is similar to buying a futures contract that

expires at time $t-i$ and then store the commodity from $t-i$ to t. Hence a similar relationship as in equation (1) also holds for two futures contracts with different time to expiration.

$$F_t^T \geq F_t^{T-i} e^{(r_t+w)(T-i)} \tag{4}$$

However, as mentioned above, this condition only says that there should be a long run relationship between the prices, the condition says nothing about whether any of the prices leads the other. If the longest contract predicts the future spot price, this should also predict the price of any shorter future contracts. On the other hand, some weighting between convenience yield and e.g. low transaction costs can lead a shorter contract to be the price leader, and as above spot prices can be price leading if the convenience yield is high enough. It is of interest to note that if equation (1) holds for all future contracts, then equation (4) follows. This is analogue to Hall, Anderson and Granger's (1992) model of interest rates.

Empirical specification

Most prices series seem to be nonstationary. Cointegration analysis is therefore the preferred tool when analyzing relationships between prices. The use of cointegration analysis and error-correction models also enables one to distinguish between short-run and long-run deviations from equilibrium indicative of price discovery and long-run deviations that account for efficiency and stability (Pizzi *et al.*, 1998). Evidence of price changes in one market generating price changes in the other market so as to bring about long-run equilibrium relationship is given as

$$s_t - \beta_0 - \beta_1 f_t = e_t \tag{5}$$

where s_t and f_t are logs of spot and futures prices at time, t; β_0 and β_1 are parameters; and e_t is the deviation from parity. If s_t and f_t are cointegrated the error term e_t will be stationary. This observation forms the basis for the Engle and Granger test for cointegration,

where e_t is tested for stationarity by performing ADF unit root tests. Engle and Granger further show that if two series, such as spot and futures prices, are cointegrated, an appropriate methodology for modeling the short-term dynamics of the system is an error-correction model (ECM). In the futures literature it has been standard practice to represent this by single equations as

$$\begin{aligned} \Delta s_t &= \alpha_1 - \theta_1 e_{t-1} \\ \Delta f_t &= \alpha_2 + \theta_2 e_{t-1} \end{aligned} \tag{6}$$

Thus, each ECM contains lagged residuals from the cointegration regression (equation (5)). A number of hypothesis is of interest in relation to equations (5) and (6). If $\beta_0 = 0, \beta_1 = 1$ the market is efficient, while if $\beta_1 = 1$ the prices are proportional and the prices are predictable from each other. If the $\theta_1 = 0$ in equation (6) the spot price is weakly exogenous for the future price and therefore leads the future price, while if $\theta_2 = 0$ the future price leads the spot price. In addition if (If $\theta_1 = \theta_2 = 0$ the spot and future prices are not cointegrated).

The Johansen test

In contrast to the Engle and Granger approach, the Johansen test is a multivariate approach. This can be represented as follows. Let X_t denote an $n \times 1$ vector, where the maintained hypothesis is that X_t follows an unrestricted vector autoregression (VAR) in the levels of the variables

$$\mathbf{X}_t = \Pi_1 \mathbf{X}_{t-1} + ... + \Pi_k \Delta \mathbf{X}_{t-k} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \tag{7}$$

where each of the Π_i is a $n \times n$ matrix of parameters, μ a constant term and ε_t are identically and independently distributed residuals with zero mean and contemporaneous covariance matrix Ω. The VAR system in (5) written in error correction form (ECM) is;

$$\Delta \mathbf{X}_t = \Gamma_1 \Delta \mathbf{X}_{t-1} + ... + \Gamma_{k-1} \Delta \mathbf{X}_{t-k+1} + \Pi \mathbf{X}_{t-k} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \qquad (8)$$

with $\Gamma_i = -I + \Pi_1 + ... + \Pi_i$, $i = 1,...,k-1$ and $\Gamma_i = -I + \Pi_1 + ... + \Pi_i$, $i = 1,...,k-1$. Hence, Π is the long-run "level solution" to (7). If X_t is a vector of $I(1)$ variables, the left-hand side and the first (k-1) elements of (8) are $I(0)$, and the kth element of (8) is a linear combination of $I(1)$ variables. Given the assumptions on the error term, the kth element must also be $I(0)$; $\Pi X_{t-k} \sim I(0)$. Hence, either X_t contains a number of cointegration vectors, or Π must be a matrix of zeros. The rank of Π, r, determines how many linear combinations of X_t are stationary. If $r=n$, the variables in levels are stationary; if $r=0$ so that $\Pi=0$, none of the linear combinations are stationary. When $0<r<n$, there exist r cointegration vectors—or r stationary linear combinations of X_t. In this case one can factor Π; $\Pi = \alpha\beta'$, where both α and β are $n \times r$ matrices, and β contains the cointegration vectors (the error correcting mechanism in the system), and α the factor loadings or adjustment parameters. Please note that the cointegration vectors are identified only up to a nonsingular transformation, so that $\alpha\beta' = \alpha PP^{-1}\beta' = AB'$ where P is an arbitrary nonsingular $r \times r$ matrix. Furthermore, the cointegration relationships give nonlinear cross equation restrictions, so the system cannot in general be estimated efficiently in single equations.

Johansen suggests two tests for the number of cointegration vectors in the system, the *maximal eigenvalue* test and the *trace* test. Both tests have the null hypothesis that there are at most r cointegration vectors. For the *maximal eigenvalue* test, the alternative hypothesis is that there are exactly $r+1$ cointegration vectors, while the alternative hypothesis in the *trace* test is that there exist more than r cointegration vectors.[iii] The Johansen procedure allows a wide range of hypothesis testing on the coefficients α and β, using likelihood ratio tests (Johansen and Juselius, 1990). When testing hypothesis with respect to basis, it is the

restrictions on parameters in the cointegration vectors β we wish to test. Information about price leadership is formally tested as exogeneity tests on the α coefficients.

It may here be worthwhile to set up a system that contains the relevant information. Let us first start with a system with two variables, a spot price s and a futures price f. Assuming that the prices are nonstationary, but cointegrated, with one lag and suppressing the error term, this can be represented as follows.

$$\begin{bmatrix} \Delta s_t \\ \Delta f_t \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ f_{t-1} \end{bmatrix} \tag{9}$$

If b_1=-b_2, the prices will be proportional and basis constant. Normally, b_1 is normalized to be 1, so that one tests whether b_2=-1 (or 1 if the term is moved to the other side of the equality operator). The a's measure the impact of changes in basis on respectively the spot and future prices. If $a_1 \neq 0$, a change in basis will be at least partly corrected by a change in the spot price, while if $a_2 \neq 0$, a change in basis will be at least partly corrected by a change in the future price. It should then be obvious that if a_1=0, there are no changes in the spot price due to changes in basis and all corrections will have to made by changes to the future price, and vice versa if a_2=0. Hence, if a_1=0 spot prices will lead futures prices, if a_2=0 future prices will lead spot prices and if $a_1 \neq a_2 \neq 0$ there will be no price leadership in this system. Both a's cannot simultaneously be zero, as there will then be no long-run relationship.

When moving to a multivariate system the main difference is that β and α is matrices (a and b in equation (9). In general, all long-run relationships influence all the variables, and since there are cross equation restrictions, information is lost if all equations are not estimated together. Although the intuition in a multivariate system is equivalent to the bivariate example above, there are also some special issues. In a system with n data series there will be r

stochastic trends and n-r cointegration vectors (Stock and Watson, 1988). If the data series are nonstationary one can at most have n-1 cointegration vectors. When all future prices are cointegrated with the spot price, they will all follow the same stochastic trend and there will be n-1 cointegration vectors in the system. This implies that β will be a $n \times n$-1matrix, and hence one can test whether all basises, or a subset of them, are constant simultaneously.

More important in our context is it that in a system with n data series and r stochastic trends there can at most be r exogenous variables (Johansen and Juselius, 1994). With the structure one expect to find in futures markets (i.e. a long-run relationship between the spot price and the futures price for all contracts), one expect to find n-1 cointegration vectors and one stochastic trend. This is in common with many other financial systems, see e.g. (Cox, Ingersoll and Ross, 1985; Hall, Anderson and Granger, 1992). Hence, there can at most be one price that leads the system. In this case it will also be valid to model the system as bivariate relationships with the exogenous variable on the right hand side. This has two important implications. First, if there are no exogenous variables in the system, the full system must be estimated. This also implies that the full system must be estimated if one is to test for exogeneity. Second, if the spot price is to be included in all relationships, the spot price must be exogenous if any prices are exogenous. In single equation specifications it must then be the right hand side variable. If it is a futures price that is exogenous, this should be the right hand side variable.

Data

The dataset used in this study consist of monthly observations of futures prices for gas oil marketed at IPE. The period covered includes the whole history of the contract, i.e. April 1981 to September 2001. Data were obtained from the Exchange's homepage (www.ipe.com).

As a proxy for the futures price, the contract closest to delivery is used. This is a common practice when cash prices are not readily available (Fama and French, 1987). Assuming risk neutral and rational actors, the future price close to delivery should represent the expected spot price when deliveries actually happen. To prevent problems with low trading volume the last day with active trading, prices quoted 5 days before are used in the analysis. With several days to the contract expires, volume should still be high enough. Futures contracts with respectively 1 month to expiration, 3 month to expiration and 6 month to expiration is used in the analysis.

Empirical Results

Before conducting any econometric analysis, the time series properties of the data must be investigated. This is done with Augmented Dickey Fuller (ADF) test. Lag length was chosen as suggested by Banerjee (1993) by starting with a generous parameterization and then removing insignificant lags. Table 1 reports the results of the ADF tests. The test indicates that while prices in levels are nonstationary, all prices are stationary in first differences. The analysis will therefore proceed under the assumption that all price series are integrated of order one.

Testing for Cointegration

Since the price series are all nonstationary and integrated of the same order, cointegration analysis is the appropriate tool to investigating the relationship between the prices. We proceed by testing for cointegration in the system containing the spot price as well as the prices for three future contracts. The lag length was again chosen to whiten the error term and as tests for autocorrelation Lagrange multiplier tests (LM) for the presence of autocorrelation up to the 12^{th} order are reported in table 2. The two tests to determine the rank of the

coefficient matrix Π, i.e. the trace and eigenvalue tests, are also reported in table 2. The maximum eigenvalue test as well as the trace test suggests that there are three cointegration vectors in the system, and accordingly only one stochastic trend. The conclusion must therefore be that spot prices and futures prices with different time to expiration are cointegrated and hence there is a long run relationship between the prices. This result is further confirmed with the results from the bivariate tests (table 3) indicating that all the prices are bilaterally cointegrated.

For the futures price to be an unbiased predictor of the spot price, the future and spot prices must be proportional (i.e $\beta_1 = 1$ in equation (5)). In the system a likelihood ratio test for this hypothesis is distributed as χ^2 (3) and give a test statistic of 2.41 with a *p*-value of 0.49. Hence the hypothesis that all the prices are proportional cannot be rejected. This result was also confirmed in bivariate relationships (test statistics in table 3). The test that the prices are proportional and in addition that β_0 (in equation (5)) are zero in all relationships was also performed with a likelihood ratio test. The test that is $\chi^2(6)$-distributed give a test statistic of 14.62 with and *p*-value of 0.02. Hence the hypothesis was rejected.

Test results so far have confirmed that spot prices and futures prices with different time to maturity follow the same stochastic trend, and move proportionally over time. However, price leadership, or causal relationship between the prices is not yet investigated. A test for exogeneity will provide such information. As discussed earlier, tests on the factor loading parameters will determine whether any of the variables can be considered as weakly exogenous in the system. The null hypothesis to be tested is that the α matrix in a particular row is containing only zeros, and is tested with a likelihood ratio test. Moreover, since there is only one stochastic trend in the system, there cannot be more then one exogenous price.

Results are reported in table 2, and indicate that weak exogeneity cannot be rejected for the futures contract with 6 months to expiration, while it is clearly rejected for the two other futures prices and the spot price. These results suggest that the futures contract with longest time to expiration is the driving factor in the price generating process and that it is this contract that binds the price series together in the long-run.

Again, the results were confirmed in the bivariate equations. Also here weak exogeneity could not be rejected for the 6 months futures contract in any of the relationship containing 6 month. In addition weak exogeneity could not be rejected for the three month contract in the bivariate equations containing three month and spot, and three month and one month. And at last, weak exogeneity cannot be rejected for the one month contract in the formulation with only one month and spot. The conclusion seem therefore to be that futures prices leads spot prices, and that futures prices with longer time to expiration leads futures contracts with shorter time to expiration, and hence it is always the longest contract that binds the price series together in the long-run.

Concluding Remarks

Tests on long run relationships and lead lag relationships between futures and spot prices are in the literature mostly carried out in a two step Engle and Granger approach. However, this approach has several shortcomings. In particular, valid tests for lead lag relationships can only be carried out in a system framework. Moreover, one cannot test hypothesis with respect to the parameters in the long-run relationship (the basis) in the Engle and Granger framework. Finally, one cannot take into account that there for most commodities are several contracts. In

this paper, it is shown that all these problems can be avoided if one use the Johansen procedure.

Our empirical results clearly demonstrate that gas oil futures and gas oil spot price form a stable long run relationship and that the prices are proportional so that the basis is constant. The results also indicate that the future price leads the spot price, and in addition, the future contracts with longer time to expiration lead futures contracts with shorter time to expiration. Hence the longest contract in our system (6 month) is the leading indicator of future spot price as well as for future futures price for shorter contracts.

Bibliography

Banerjee, A., Dolado, J., Galbraith, J. W. & Hendry, D. F. (1993). Co-integration, error correction, and the econometric analysis of non-stationary data, Oxford: Oxford University Press.

Breeden, D. T. (1980). Consumption risks in future markets. Journal of Finance, 35:503-520.

Brennan, M. J. (1958). The supply of storage. American Economic Review, 48:50-72.

Chan. K. (1992). A further analysis of the lead-lag relationship between the cash market and stock index futures market. Review of Financial Studies, 5, 123-152.

Cox, J., Ingersoll, E., & Ross, S. (1985). A theory of the terms structure of interest rates. Econometrica, 53, 385-407.

Cootner, P. H. (1960). Returns to speculators: Telser vs. Keynes. Journal of Political Economy, 68, 396-404.

Dusak, K. (1973). Futures trading and investor returns: An investigation of commodity market risk premiums. Journal of Political Economy, 81, 1387-1406.

Engle, R. F., & Granger C. W. J. (1987). Co-integration and error correction, representation estimation and testing. Econometrica, 55, 251-276.

Fama, E. F., & French, K. R. (1987). Commodity futures prices: some evidence on forecast power, premiums, and the theory of storage. Journal of Business, 60(1), 55-73.

Garbade, K. D., & Silber, W. L. (1983). Price movements and price discovery in futures and cash markets. Review of Economics and Statistics, 65, 289-297.

Girma, P. B., & Paulson, A. S. (1999). Risk arbitrage opportunities in petroleum futures spreads. The Journal of Futures Markets, 19(8), 931-955.

Gülen, S. G. (1998). Efficiency in the crude oil futures market. Journal of Energy Finance & Development, 3(1), 13-21.

Haigh, M. S. (2000). Cointegration, unbiased expectations, and forecasting in the BIFFEX freight futures market. The Journal of Futures Markets, 20(6), 545-571.

Hall, A.D., Anderson, H. M., & Granger, C. W. J. (1992). A cointegration analysis of treasury bill yields. The Review of Economics and Statistics, 74(1), 116-126.

Hazuka, T. B. (1984). Consumption betas and backwardation in commodity markets. Journal of Finance, 39, 647-655.

Herbst, A. F., McCormack, J. P., & Koch, T. W. (1987). Investigation of a lead-lag relationship between spot stock indices and their futures contracts. The Journal of Future Markets, 7, 373-381.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration - with applications to the demand for money. Oxford Bulletin of Economics and Statistics, 52, 169-210.

Johansen, S., & Juselius., K. (1994). Identification of the Long-Run and the Short-Run structure: an application to the IS-LM model. Journal of Econometrics, 63, 7-36.

Johansen, S. (1991) Estimation and hypothesis-testing of cointegration vectors in gaussian vector autoregressive models. Econometrica, 59(6), 1551-1580.

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

Kaldor, N., (1939). Speculation and economic stability. Review of Economic Studies, 7, 1-27.

Kawaller, I., Koch, P., & Koch, T. (1987). The temporal price relationship between S&P 500 futures and the S&P 500 index. Journal of Finance, 42, 1309-1329.

Kellard, N., Newbold, P., Rayner T., & Ennew C. (1999). The relative efficiency of commodity futures markets. The Journal of Futures Markets, 19(4), 413-432.

Moosa, I. A. (1996). An econometric model of price determination in the crude oil futures markets, in M. McAleer, P. Miller & Leong, K. (Eds), Proceedings of the Econometric Society Australian meeting (vol 3, pp. 372-402), Perth: University of Western Australia.

Moosa, I. A., & Al-loughani, N. E. (1994). Unbiasedness and time-varying risk premia in the crude-oil futures market. Energy Economics, 16(2), 99-105.

Newberry, D. M. (1992). Futures markets: hedging and speculation, in P. Newman and M. Milgate, & J. Eatwell (Eds), The new palgrave dictionary of money and finance. 2, 207-210.

Pizzi, M. A., Economopoulos, A. J. & O'Neill, H. M. (1998). An examination of the relationship between stock index cash and futures markets: A cointegration approach. The Journal of Future Markets, 18, 297-305.

Quan, J. (1992). Two step testing procedure for price discovery role of futures prices. The Journal of Futures Markets, 12(2), 139-149.

Schwartz, T., & Szakmary. A. (1994). Price discovery in petroleum markets: Price discovery, cointegration and the time interval of analysis. The Journal of Futures Markets, 9, 147-167.

Silvapulle, P., & Moosa, I. A. (1999). The relationship between spot and futures prices: evidence form the crude oil market. The Journal of Futures Markets, 19(2), 175-193.

Stock, J. H. & Watson, M. W. Testing for common trends. Journal of the American Statistical Association, 83, 1097-1107.

Telser, L.G. (1958). Futures trading and the storage of cotton and wheat. Journal of Political Economy, 66, 233-255.

Working, H. (1948). Theory of the inverse carrying charge in futures markets. Journal of Farm Economics, 30, 1-28.

Working, H. (1949). The theory of the price of storage. American Economic Review, 39, 1254-1262.

Table 1. Dickey-Fuller tests.

	Price levels	First differences
Spot	-2.44	-11.74**
1 month	-2.36	-11.48**
3 months	-2.45	-11.02**
6 months	-2.63	-10.89**

** indicates significance at a 1 percent level. Critical values at the five percent level is -2.87 and at the one percent level -3.46.

Table 2. Multivariate Johansen Tests.

Variables	H_0:rank = p	Max test [a)]	Trace test[a)]	LM[b)]	Exogeneity[c)]
Spot	P= =0	74.2**	152.3**	1.72 (0.06)	10.10 (0.01)*
1 month	P<=1	50.63**	78.07**	1.59 (0.09)	12.37 (0.00)**
3 months	P<=2	20.18*	27.43**	1.02 (0.43)	10.39 (0.02)*
6 months	P<=3	7.25	7.251	0.78 (0.67)	3.35 (0.34)

[a)] Critical values for the cointegration test can be found in Johansen and Juselius (1990).

[b)] LM is a Lagrange multiplier test against autocorrelation up to twelve lags for the equation for the associated variable in column 1. The test are distributed as $F(12,208)$, p-values in parentheses

[c)] The test is distributed as $\chi^2(3)$.which have a critical value at 7.82 at a 5% level. ** indicates statistical significance at 1% while * indicates statistical significance at 5% level. p-values in parentheses.

Table 3. Bivariate Johansen Tests.

Variables	H_0:rank = p	Max test[a)]	Trace test[a)]	Const basis[b)]	LM[c)]	Exogeneity[d)]
Spot	P==0	43.43**	52.48**	3.35	1.25 (0.25)	9.95 (0.00)**
1 Month	P<=1	9.05	9.046	(0.07)	1.38 (0.18)	1.62 (0.20)
Spot	P==0	39.44**	47.47**	2.65	1.47 (0.14)	15.14 (0.00)**
3 months	P<=1	8.026	8.026	(0.10)	0.93 (0.52)	0.70 (0.40)
Spot	P==0	33.04**	40.63**	1.22	1.63 (0.08)	16.23 (0.00)**
6 months	P<=1	7.588	7.588	(0.27)	0.60 (0.84)	0.78 (0.38)
1 month	P==0	42.73**	50.19**	2.45	1.26 (0.25)	12.04 (0.00)**
3 months	P<=1	7.462	7.462	(0.12)	0.70 (0.75)	3.25 (0.07)
1 month	P==0	42.85**	50.36**	1.41	1.08 (0.37)	16.33 (0.00)**
6 months	P<=1	7.504	7.504	(0.24)	0.54 (0.89)	2.05 (0.15)
3 months	P==0	48.81**	56.5**	0.81	0.32 (0.99)	13.60 (0.00)**
6 months	P<=1	7.69	7.69	(0.37)	0.34 (0.98)	2.76 (0.10)

[a)] Critical values for the cointegration test can be found in Johansen and Juselius (1990).

[b)] The test is distributed as $\chi^2(1)$, which have a critical value at 3.84 at a 5% level.

[c)] LM is a Lagrange multiplier test against autocorrelation up to twelve lags for the associated variable in column 1. The test are distributed as $F(12,221)$, p-values in parentheses

[d)] The test is distributed as $\chi^2(1)$, which have a critical value at 3.84 at a 5% level, ** indicates statistical significance at 1% while * indicates statistical significance at 5% level, p-values in parentheses.

Footnotes

[i] If the data series are stationary, one can test for the existence of a long run relationship with traditional statistical tools.

[ii] Silvapulle and Moosa (1999) present an overview of the different argument used.

[iii] Both tests have non-standard distributions. Critical values for these tests have been tabulated by Johansen and Juselius (1990).