PRICE DISCOVERY IN PETROLEUM MARKETS: ARBITRAGE, COINTEGRATION, AND THE TIME INTERVAL OF ANALYSIS

THOMAS V. SCHWARZ
ANDREW C. SZAKMARY

INTRODUCTION

A recent article in this journal [Quan (1992, p. 144)] concludes that, "...the spot [crude oil] market always leads the futures market and the crude oil futures market does not play an important role in price discovery." This is a startling statement given the tremendous growth and success of the energy futures contracts traded at the New York

Please address all correspondence to: Thomas V. Schwarz, Department of Finance, Southern Illinois University at Carbondale, Carbondale, IL 62901. (618)453-1423; Fax: (618)453-7961. INTERNET: GA3681@SIUCVMB.siu.edu
We gratefully acknowledge the help and data provided by Joe Link of *Platt's Oilgram Price Report* and Ben Windham and Cynthia Streif of *Citgo Petroleum Corporation* as well as the helpful comments of Shubash Sharma and two anonymous referees. Computational assistance by Seong-Hoon Yu is greatly appreciated. We remain solely responsible for any errors.

- *Thomas V. Schwarz is an Assistant Professor of Finance at Southern Illinois University.*
- *Andrew C. Szakmary is an Assistant Professor of Finance at Southern Illinois University*

The Journal of Futures Markets, Vol. 14, No. 2, 147–167 (1994)

CCC 0270-7314/94/020147-21

Mercantile Exchange (NYMEX).[1] Why do these futures markets continue to flourish if they do not provide one of the essential tenants for their existence? Are these markets pure satellites serving only as a tool for risk transference, if this is even at all possible? Why should petroleum markets be unique among all other assets in not demonstrating futures price leadership?[2] The magnitude of these questions necessitates a further examination of Quan's study and energy futures markets in general.

This article extends previous petroleum research in several significant ways. First, the time interval between price observations is reduced to a single day. Previous studies on non-petroleum contracts have shown that futures price leadership is often short, sometimes running only minutes ahead of the cash market. A lengthy interval of analysis, as in the use of monthly observations in Quan, may fail to capture the short-term dynamics of these price series.

Second, daily spot market data for the deliverable commodity is used in comparison with futures prices. This spot data is measured as of the close of futures trading on the NYMEX. Consequently, simultaneous spot and futures prices are used. Other studies, such as Serletis and Banack (1990) and Cho and McDougall (1990), use the nearby futures contract as a surrogate for the spot market price in an analysis with the first deferred futures price series. While this approach is not uncommon in the literature, it is suboptimal for investigating the price discovery role of futures markets when spot data is available.[3]

Third, the analysis includes two other actively traded energy futures contracts in addition to crude oil; namely, heating oil and unleaded gasoline. These latter contracts represent refined products and have a

[1]The three major petroleum futures contracts traded on the New York Mercantile Exchange have set volume records in each year of their existence. Currently, 1992 (through June) average daily volume levels are 88,296; 29,517; and 28,554 contracts for light sweet crude oil futures, heating oil futures, and unleaded gasoline futures, respectively. They rank as the 3rd, 13th, and 12th most active futures contracts in the U.S. In addition, there are active options markets on each of these futures contracts and the exchange continues to introduce additional petroleum products such as sour crude futures and futures options, natural gas futures, and propane futures.

[2]A short list of recent studies that have found futures to dominate in price leadership is representative of many commodities: live cattle futures [Bessler and Covey (1991)], live hog futures [Schroeder and Goodwin (1991)]; barley, oats, and canola futures [Khoury and Yourougou (1991)]; and stock index futures [Stoll and Whaley (1990); Schwarz and Laatsch (1991)]. See the reference list of these sources for a more extended listing.

[3]For example, the conclusions of Serletis and Banack (1990) are misleading when they state, "we showed that the current spot price dominates the current futures price in explaining movements in the future spot price" (p. 378). What they really analyzed was the relationship between the nearby and first deferred futures contracts. As a consequence, their analysis ignores the relative transactions costs differential between the actual spot and futures markets.

relatively more seasonal fluctuation in their supply and demand components [see Cho and McDougall (1990)]. This additional fluctuation in underlying fundamentals provides an enhanced opportunity to observe and compare price leadership between spot and futures. Finally, several methodological improvements relative to earlier studies are employed, including testing for price leadership within a cointegration framework and using a single equation approach to estimate the Garbade and Silber (1983) model. The latter allows a formal test of hypotheses concerning price leadership that in previous work is examined only in an ad hoc manner.

The following section explains how the cointegration framework of Engle and Granger (1987) can be theoretically expanded under the Garbade and Silber (1983) model when testing for the relationship between spot and futures price series. Here, it is shown that the cointegration of spot and futures markets is a function of the elasticity of arbitrage capital between these markets. Consequently, if these markets are as actively arbitraged as other spot–futures markets, they should be fully cointegrated. Furthermore, price leadership in these markets is shown to be a function of relative market size, which itself may be affected by asymmetric transactions costs between the two markets.

ARBITRAGE AND COINTEGRATION

Garbade and Silber (1983), hereafter GS, presented a model of simultaneous price dynamics for spot and futures prices.[4] In perfect markets (i.e., without restrictions on arbitrage services), spot and futures prices are in partial equilibrium if:

$$F_t = S_t(1 + r)^{T-t} \tag{1a}$$

This equation says that the futures price at time t, F_t, is equal to the spot, S_t, (cash) price plus a premium due to the deferred payment on a futures contract (equal to the financing cost at a rate of return, r, until contract maturity at time T).[5] Since the financing cost of the spot commodity erodes in a fairly predictable manner with the passage of time, its effect is extracted from the futures price to create a "spot equivalent" futures

[4]This section presents an abridged version of the Garbade and Silber model and does not fully capture the richness found in the original work.
[5]The original version is presented in natural log form such that F_t = ln (futures price) and similarly for the cash market. To make a comparison to standard cointegration models, the model here is presented in absolute price form. Tests are conducted in both levels and natural log form.

price, F_t', resulting in a restatement of (1a) as:[6]

$$F_t' = S_t \tag{1b}$$

where

$$F_t' = F_t/(1 + r)^{T-t}$$

It is important to note that expectations of future spot prices do not appear in the equation because they affect the current as well as the futures price. Under perfect market assumptions, eq. (1) must hold because any violation will mobilize an infinite supply of arbitrage activity in simultaneously buying the cheap and selling the expensive asset, thereby locking in a riskless profit.

GS extended this pricing model to the more realistic case of imperfect markets where the supply of arbitrage services is finite.[7] That extension yielded separate pricing equations for futures and cash markets which are (potentially) linked by arbitrage activity. If there is no arbitrage, it was shown that the two market prices will be independent; whereas, with an infinite supply of arbitrage services, the two equations converge to eq. (1). This development highlights the role of arbitrage services as the link between spot and futures markets.[8]

A final extension of the model from a static to a dynamic equilibrium process (by assuming a random walk process for both cash and futures reservation prices) yields the following model of simultaneous price dynamics:

$$\begin{bmatrix} S_t \\ F_t' \end{bmatrix} = \begin{bmatrix} 1 - \beta_s & \beta_s \\ \beta_f & 1 - \beta_f \end{bmatrix} \begin{bmatrix} S_{t-1} \\ F_{t-1}' \end{bmatrix} + \begin{bmatrix} U_t^s \\ U_t^f \end{bmatrix} \tag{2}$$

where the specifically defined parameters, β_s and β_f, are functions of the elasticity of supply of arbitrage services, the number of participants

[6]Cho and McDougall (1990) used the same adjustment in their examination of storage cost for energy futures. Our subsequent empirical analysis was conducted for both F_t and F_t'. Similar results were obtained in both cases.

[7]For energy products, some restrictions to arbitrage services include: (1) arbitrage risks (e.g., fluctuating financing rates and convenience values, execution risks), (2) a limited availability of storage capacity, transportation means, and arbitrage capital, (3) heterogeneity in the possible deliverable grades and their location in the spot market, (4) restrictions in the ability to short sell, and (5) a stochastic convenience yield from inventory holdings.

[8]Arbitrage services is defined in a general sense to include any activity that would allow for either the spot or futures market to serve as a close substitute for the other. Since the spot products evaluated in this study are deliverable into the futures contract, the delivery mechanism itself can be viewed as a source of arbitrage pressure.

in each market, and the elasticity of demand for the cash and futures instruments.

An overview and summary of the cointegration model of Engle and Granger (1987), hereafter EG, is presented next. In their model, two nonstationary time series are said to be cointegrated if a linear combination of them is found to be stationary. In essence, in cointegrated time series, a long-run component (leading to nonstationarity in the original series) that is common to both time series is canceled out by a linear combination of the two. Formally, the spot and spot-equivalent futures price series will be said to be cointegrated if the residual, Γ_t, in eq. (3) is found to be stationary, where τ is the cointegrating coefficient.

$$F_t^l = \mu + \tau S_t + \Gamma_t \tag{3}$$

The relationship between cointegration and price determination in spot and futures markets can now be highlighted. In the perfect market of eq. (1) where there is an infinite supply of arbitrage services, spot-equivalent futures and spot prices *must* be perfectly cointegrated; the cointegrating coefficient will equal 1, and the residual Γ_t will equal 0 for all t.[9] If less than perfect arbitrage is available, the two price series will move according to eq. (2). In theory, arbitrage activity should prevent the two series from drifting too far apart, and they will, therefore, be cointegrated.[10] Only in the special case where the supply of arbitrage services is zero (for which GS show that the coefficients $\beta_s = \beta_f = 0$ and, therefore, both cash and futures are independent random walks) will spot and spot-equivalent futures prices *not* be cointegrated.

In summary, the formal cointegration test of EG, when applied to spot and futures prices on the same product, is primarily a test of the level of arbitrage activity between the two markets. Since the delivery feature of a futures contract is such a strong (an ultimately absolute) force in arbitraging these two series, it is very unlikely that the deliverable spot and its nearby futures would not be cointegrated. In fact, the GS model shows that the price discovery and hedging functions *require* the presence of cointegrated series.

[9] The constant μ may represent any systematic differences in spot and futures pricing, e.g., quality or location differential, etc. Interestingly, tests for cointegration, which involve examining the stationarity of the residual term from the cointegrating regression, cannot be implemented if two series are perfectly tied together in the short and long run such that the residual is always zero. Fortunately, this result is very unlikely to occur in practice.

[10] In practice, if arbitrage activity between spot and futures is present but very weak, it is possible that with a finite dataset and limited observations, the two series will be found not to be cointegrated. The null hypothesis in these tests is non-cointegration; with weak arbitrage activity the tests may lack sufficient power to reject the null.

DATA

Daily closing futures prices for crude oil, heating oil, and unleaded gasoline traded at the NYMEX are used as the base for constructing time intervals of various lengths and are obtained from Tick Data, Inc. The nearby futures is used until the day preceding its last day of trading when data for the next deferred contract is used. These rollover periods are controlled for in the daily analysis by eliminating any price change due to this shift and by the interest adjustment in calculating futures prices.[11] The period of analysis extends from January 1, 1984, to May 15, 1991, for crude oil and heating oil futures with unleaded gasoline beginning on January 1, 1985.[12]

The daily spot price data for the par deliverable grade of West Texas Intermediate crude and New York unleaded gasoline (barge) is obtained from *Platt's Oilgram Price Report,* the industry leader in price reporting. Platt's prices represent an assessment made at the close of NYMEX trading based on transactions and market information from sources deemed reliable and active in the spot markets. Spot price data for heating oil is from Tick Data, Inc. This price series matches the cash prices reported daily in the *Wall Street Journal.*

METHODOLOGY

Unit Root Tests

Proper interpretation of cointegration models in the context of this study requires that spot and adjusted futures prices contain a single unit root. Two procedures are implemented to ascertain the unit root properties of this data. The first of these—the Augmented Dickey Fuller (ADF) unit root tests—are conducted by estimating the following OLS regressions

[11] Ma, Mercer, and Walker (1992) report that biases can be generated from the selection of different methods to rollover futures contracts. Two decisions to be made include the selection of a rollover date and the adjustment of price at the time of rollover [see eq. (1b)]. In the calculation of F', the predictable interest effect at the date of rollover is removed. In addition, since there is considerable distance between the termination of trading and delivery, rollover on the day preceding the last day of trade is chosen. The last day of trade for light crude oil futures is the third business day prior to the 25th of the month preceding the delivery month. However, the actual delivery cannot be initiated until the delivery month and there is great uncertainty for both long and short since the oil dispatching schedule is maintained by independent pipeline companies [see Ma (1989)]. For heating oil and unleaded gasoline futures, the last day of trading is the last business day in the month prior to delivery. Uniquely, the long determines a 5-day delivery window in which the short must make delivery. Ma (1989) finds that the optimal delivery day is generally the first day of this window.

[12] Light sweet crude oil, heating oil #2, and unleaded gasoline futures first began trading on the NYMEX on 3/30/83, 11/23/78, and 12/3/84, respectively.

for each series:

$$\Delta X_t = \zeta_1 + \eta_1 X_{t-1} + \sum_{i=1}^{N} \lambda 1_i \Delta X_{t-i} + e1_t \qquad (4a)$$

and

$$\Delta^2 X_t = \zeta_2 + \eta_2 \Delta X_{t-i} + \sum_{i=1}^{N} \lambda 2_i \Delta^2 X_{t-i} + e2_t \qquad (4b)$$

where X is the series investigated, $\Delta X_t = X_t - X_{t-1}$, and $\Delta^2 X_t = \Delta X_t - \Delta X_{t-1}$. The lag length N is chosen just large enough to eliminate significant autocorrelation in the residuals. The null hypothesis in model (4a) is that X_t has one unit root; the null is rejected in favor of zero unit roots (i.e., stationarity) if the t-statistic on the coefficient η_1 exceeds critical values tabulated in Fuller (1976). Model (4b) is used to determine if the series has multiple unit roots; the null of two unit roots is rejected (in favor of the alternative of one unit root) if the t-statistic on η_2 exceeds critical values.

Unit root tests have been criticized on the grounds that failure to reject the null of a unit root does not conclusively show that a unit root exists. It is well known that these tests have low power to distinguish between a unit root and weakly-stationary alternatives. Consequently, we also implement a stationarity test described in Kwiatkowski, Phillips, Schmidt, and Shin (1992). In this test, the null hypothesis is that a series is stationary, with the alternative being that the series has a unit root. The test statistic is calculated as

$$\hat{\eta}_u = T^{-2} \sum_{t=1}^{T} S_t^2 / s^2(L) \qquad (5a)$$

where T is the number of observations, S_t is the cumulative sum of the residuals (e_t) from a regression of the series being tested on a constant (i.e., $S_t = \sum_{i=1}^{t} e_i$, $t = 1, 2, \ldots, T$), and where

$$s^2(L) = T^{-1} \sum_{t=1}^{T} e_t^2 + 2T^{-1} \sum_{s=1}^{L} (1 - s/(L+1)) \sum_{t=s+1}^{T} e_t e_{t-s} \qquad (5b)$$

If the calculated test statistic exceeds the critical values in Kwiatkowski et al. (1992, Table 1, p. 166), the null hypothesis of stationarity can be rejected in favor of the unit root alternative. The Bartlett lag window used to calculate $S^2(L)$ is arbitrary; higher values of L increase reliability when the e_t's are autocorrelated, but reduce the power of the test. Herein, $L = 24$ because this is the longest L evaluated by Kwiatkowski et al., and because they find that the power of the test is largely undiminished with this value of L when the sample size is large (i.e., greater than 500).[13] The stationarity test is implemented for both levels and first differences of cash and adjusted futures prices.

Cointegration Modeling

Engle and Granger (1987) cointegration tests are conducted using spot and adjusted futures prices from each market. These tests are implemented by performing ADF unit root tests on the residuals from the cointegrating regression (3). The model estimated is

$$\Delta\Gamma_t = \psi\Gamma_{t-1} + \sum_{i=1}^{N} W_i \Delta\Gamma_{t-i} + e_{t_1} \tag{6}$$

where Γ_t is the series of residuals from the cointegrating regression, eq. (3). Note that if F_t' and S_t each contain a single unit root, Γ_t can have at most one root, rendering a test for multiple roots (as in 4b) unnecessary. Also, (6) contains no constant term because Γ_t, by construction, has zero mean. Finally, note that the critical values for the t-statistic of ψ are somewhat higher than in Fuller (1976) because Γ_t is constructed from two variables; the appropriate critical values are provided in Engle and Yoo (1987).

If two series, such as spot and futures prices, are cointegrated, Engle and Granger (1987) show that an appropriate methodology for modeling the short-term dynamics of the system is an error-correction model (ECM). The following ECMs are estimated:

$$\Delta S_t = \alpha + \beta\Gamma_{t-1} + \sum_{i=1}^{N} \gamma_i \Delta F_{t-i}' + \sum_{i=1}^{N} \phi_i \Delta S_{t-i} + e_t \tag{7a}$$

$$\Delta F_t' = \alpha + \beta\Gamma_{t-1} + \sum_{j=1}^{M} \gamma_j \Delta S_{t-j} + \sum_{j=1}^{M} \phi_j \Delta F_{t-j}' + e_t \tag{7b}$$

Thus, each ECM contains the lagged residual from the cointegrating regression (Γ_{t-1}), as well as lagged first differences in spot and adjusted futures prices. The optimum lag lengths N and M are chosen by minimizing the Schwarz (1978) criterion for each ECM. Additionally,

[13]Given the sample sizes (1600 or 1849, depending on the product), $L = 24$ corresponds to the $L12$ = integer $[12(T/100)^{1/4}]$ window evaluated in Kwiatkowski et al. Tests using $L = 8$ (corresponding to the $L4$ window of Kwiatkowski et al.) were also run which provided similar conclusions regarding the non-stationarity of the tested series.

the lag lengths for ΔF_t and ΔS_t within each ECM are constrained to be equal.[14] Note that since all variables in an ECM are stationary, standard hypothesis tests are applicable to the estimated coefficients. If necessary, covariance matrices can be estimated that are robust in the presence of heteroskedasticity and autocorrelation in the residuals, using the technique of Newey and West (1987).

The primary purpose in estimating the ECMs is to implement price leadership tests between S_t and $F_t^{\prime}$ in each market. Granger (1988) showed that tests of causality between cointegrated variables must be conducted in an error-correction framework, because standard tests of causality overlook the reversion-to-equilibrium channel of causality represented by $\beta\Gamma_{t-1}$ in (7a) and (7b). Causality tests in the ECM framework involve testing exclusion restrictions on the coefficients β and all γ_i in (7a) and (7b). If these coefficients are jointly insignificant in (7a), then futures do not Granger-cause spot prices, and there is no evidence of price leadership from futures to spot. Similarly, if these same coefficients are jointly insignificant in (7b), then there is no evidence of price leadership from spot to futures. If futures and spot prices are cointegrated, there *must* be Granger causality in at least one direction.

Garbade and Silber Model

The usual procedure for implementing the GS model is to rearrange eq. (2) to provide the empirically testable form below:

$$(S_t - S_{t-1}) = \alpha_s + \beta_s(F_{t-1}^{\prime} - S_{t-1}) + U_t^s \tag{8a}$$

$$(F_t^{\prime} - F_{t-1}^{\prime}) = \alpha_f + \beta_f(S_{t-1} - F_{t-1}^{\prime}) + U_t^f \tag{8b}$$

where both spot and adjusted futures prices are measured in logarithmic form. Each equation is then estimated using ordinary least squares. In these regressions, the change in each market price is a function of the mispricing $(F_{t-1}^{\prime} - S_{t-1})$ that results from market imperfections. That market which shows the greatest price discovery is the one that does

[14]The Schwarz criterion is used and the lag lengths are constrained to be equal to minimize problems associated with data-snooping. Clearly, a two-step procedure in which one first lets the data determine the "best" model and then performs joint significance tests on the included variables is biased in favor of finding significance even where it doesn't truly exist. The magnitude of this problem will be a function of the size of the models and the number of models considered in the first stage. The Schwarz criterion imposes a larger penalty on additional variables and results in smaller models than the Akaike criterion. Constraining the lags of ΔS_t and ΔF_t to be equal within each ECM results in a large reduction in the number of models considered.

not follow but rather initiates the market mispricing; that is, that market with the lowest coefficient, β_s or β_f. The measure $\theta = \beta_s/(\beta_s + \beta_f)$, which is theoretically bound between zero and one, can be used to show relative price leadership. Values of θ greater than 0.5 show relatively greater futures price leadership and vice versa. Equation (8) is then solved for the mispricing variable $(F'_{t-1} - S_{t-1})$ as below:

$$(F'_t - S_t) = \alpha + \delta(F'_{t-1} - S_{t-1}) + e_t \tag{9}$$

The coefficient δ, which must equal $1 - \beta_f - \beta_s$, indicates the rate of convergence in the mispricing series and is inversely related to the supply of arbitrage services. Consequently, δ will be inversely related to the presence of cointegration in the two time series. Values of δ very close to one are unlikely to be associated with cointegrated series; whereas, lower values will be indicative of cointegration.

The usual implementation of the GS model described above suffers from the shortcoming that no formal statistical test for θ can be conducted. This is a serious problem given that the direction of price leadership is the central issue in this study. To allow a formal test of θ with respect to some value (such as 0.5), a single equation approach is implemented using "stacked" cross-section/time series data and dummy variables. As discussed below, the methodology yields coefficient estimates which are identical to those obtained by estimating regressions (8a), (8b), and (9). In addition, it is now possible to conduct statistical tests for θ.

A "stacked" dataset with $2T$ observations, where T is the number of observations available for both spot and futures prices is constructed so as to estimate the following regression:

$$\begin{aligned}\Delta P_t = &\alpha_s DS_t + \beta_s[DS_t * (F'_{t-1} - S_{t-1})] + \alpha_f DF_t \\ &+ \beta_f[DF_t * (S_{t-1} - F'_{t-1})] + \mu_t\end{aligned} \tag{10}$$

where DS_t and DF_t are 0, 1 dummy variables denoting spot and futures prices. Thus, $DS_t = 1$ if observation t represents a spot price and 0 otherwise; $DF_t = 1$ if observation t represents a futures price, and 0 otherwise. ΔP_t can be formally defined as $DF_t(F'_t - F'_{t-1}) + DS_t(S_t - S_{t-1})$, where either DF_t or DS_t (but not both) equals one. Thus, ΔP_t is the change in price, and does not distinguish between spot and futures price changes. Clearly, for all observations for which $DF_t = 0$ (i.e., for the spot price changes), eq. (10) reduces to (8a). Similarly, for all observations where $DS_t = 0$ eq. (10) reduces to (8b). Thus, the coefficient estimates for α_s, β_s, α_f, and β_f obtained from estimating (10) will be identical to those obtained by estimating (8a) and (8b). However, now that β_s and β_f are estimated in the same regression, it is possible to conduct formal statistical tests of any linear combination involving these coefficients. The hypothesis $\theta = \beta_s/(\beta_s + \beta_f) = 0.5$ can be simplified algebraically to the linear restriction, $\beta_s - \beta_f = 0$. Similarly, the hypothesis $\delta = 1 - \beta_f - \beta_s = 1$ can be tested as $\beta_s + \beta_f = 0$.

Another advantage of estimating the "stacked" regression (10) in place of regressions (8a), (8b), and (9) is the ease with which possible heteroskedasticity and/or autocorrelation in the residuals can be handled. As in the case of the ECM's, the stacked regressions can be estimated with the Newey and West (1987) adjustment to the covariance matrix, rendering the estimates of the standard errors robust to heteroskedasticity and autocorrelation.

EMPIRICAL RESULTS

Stationarity Tests

Table I presents the results for the ADF unit root tests for spot and futures for all three energy products. In each case, save heating oil spot prices at the 10% level, the t-statistic on the η parameter is insufficient to reject the null hypothesis of a unit root in levels. Panel B results soundly reject the presence of a second unit root suggesting that all first differences are stationary.

The Kwiatkowski et al. stationarity tests are reported in Table II. The results in Panel A show that the null hypothesis of stationarity of price levels is rejected (in favor of a unit root) for both spot

TABLE I
Augmented Dickey Fuller Unit Root Tests[a]

Panel A: Test for Single Unit Root

$$\Delta X_t = \zeta_1 + \eta_1 X_{t-1} + \sum_{i=1}^{N} \lambda_1 i \Delta X_{t-i} + e_t$$

	Crude Oil[b]	*Heating Oil*[b]	*Unleaded Gasoline*[c]
Spot	−2.124	−2.668[e]	−1.988
Futures[d]	−1.711	−1.955	−0.514

(continued)

TABLE I (continued)

	Panel B: Test for Two Unit Roots $\Delta^2 X_t = \zeta_2 + \eta_2 \Delta X_{t-1} + \sum_{i=1}^{N} \lambda_2 i \Delta^2 X_{t-i} + e_t$		
	Crude Oil[b]	*Heating Oil*[b]	*Unleaded Gasoline*[c]
Cash	−11.745[f]	−11.801[f]	−14.097[f]
Futures[d]	−10.672[f]	−11.023[f]	−15.551[f]

[a] $\Delta X_t = X_t - X_{t-1}$ and $\Delta^2 X_t = \Delta X_t - \Delta X_{t-1}$. The numbers reported are the t-statistics for the hypothesis that $\eta = 0$ in the regressions listed. Critical values are given in Fuller (1976). Lag length N is chosen such that the Q-statistic at 24 lags indicates absence of autocorrelation in the residuals. All regressions are estimated using OLS.
[b] Estimation period is Jan. 1, 1984–May 15, 1991 (1762 observations after elimination of days when futures roll over). The number of observations used in the regressions is somewhat lower to allow for lagged values.
[c] Estimation period is Jan. 1, 1985–May 15, 1991 (1524 observations after elimination of days when futures roll over).
[d] Futures prices are adjusted for interest carrying cost through the last day of trading.
[e] Indicates significance at the 10% level.
[f] Indicates significance at the 1% level.

and futures for all three products. Stationarity is rejected at the 1% level for crude oil and heating oil, and at the 5% level for unleaded gasoline. When the test is repeated using first differences (Panel B), stationarity is not rejected for any series. Thus, the results of the Kwiatkowski tests strongly indicate the presence of a single unit root in every spot and futures series, and reinforce the ADF tests in Table I. These findings accord with a priori expectations that asset prices in a reasonably efficient market should *not* be stationary—and hence predictable. The results also illustrate that spot and futures prices for each product show similar temporal properties. This result is as expected since it is a prerequisite for hedging, one of the primary functions of these markets.

Cointegration Tests

Table III reports the results for Engle and Granger (1987) cointegration tests between the spot and futures markets of each product. In each of the three energy products, futures prices are strongly cointegrated with their deliverable spot price. In particular, the residuals from the cointegrating regressions are highly stationary (t -test on ψ highly significant). These results suggest that the spot and futures markets for each product are subject to the same nonstationary properties.

TABLE II
Stationarity Tests[a,b]

Panel A: Test for Stationarity of Price Levels

	Crude Oil	Heating Oil	Unleaded Gasoline
Spot	1.827[d]	1.077[d]	0.466[e]
Futures[c]	1.843[d]	1.071[d]	0.460[e]

Panel B: Test for Stationarity of First Differences in Prices

	Crude Oil	Heating Oil	Unleaded Gasoline
Spot	0.095	0.000	0.000
Futures[c]	0.081	0.000	0.000

[a]The test statistic is $n_u = T^{-2}\sum S_t^2/s^2(L)$, where

$$S_t = \sum_{i=1}^{t} e_i, \quad t = 1,2,\ldots,T; \text{ and}$$

$$s^2(L) = T^{-1}\sum_{t=1}^{T} e_t^2 + 2T^{-1}\sum_{s=1}^{L}(1 - S/(L+1)) \sum_{t=S+1}^{T} e_t e_{t-s}$$

The e_t's are the residuals from a regression of the series being tested on a constant. Critical values are given in Kwiatkowski et al. (1992, Table 1, p. 166). The reported test statistics are computed using the lag length $L = 24$.
[b]Estimation period is Jan. 1, 1984–May 15, 1991 (1849 observations) for crude oil and heating oil and Jan. 1, 1985–May 15, 1991 (1600 observations) for unleaded gasoline.
[c]Futures prices are adjusted for interest carrying cost through the last day of trading.
[d]Indicates significance at the 1% level.
[e]Indicates significance at the 5% level.

Consequently, the futures market is likely to serve as a viable hedge for the spot commodity.[15]

Table III also reports that the cointegrating parameter in each case is significantly less than one, suggesting that the futures price is less volatile than the spot for all three energy markets. It must be noted

[15]To examine the sensitivity of the results to model specification, the Engle and Granger tests are repeated using the reverse cointegrating regression $S_t = \mu + \tau F_t' + \Gamma_t$. These results are very similar to those reported in Table III. In addition, on the advice of an anonymous referee, Johansen tests for cointegration of spot and futures prices were implemented for each product. This procedure is described in Johansen (1988, 1991), Hall (1989), and Baillie and Bollerslev (1989). In theory the Johansen test is superior to the Engle and Granger approach since it treats variables as jointly endogenous and tests cointegration in a multivariate framework. In practice, however, it has been shown to be extremely sensitive to the sample period over which it is estimated and, thus, to produce fragile inferences [see Sephton and Larsen (1991)]. In any case, for the data and sample period of this study, the Johansen tests strongly indicated the presence of one cointegrating vector for all three pairs of spot and futures prices, a result which is identical to the findings using the Engle and Granger approach. (The Johansen test results, not reported, are available from the authors on request.)

TABLE III
Cointegration Tests[a]

Cointegrating Regression: $F_t^{\prime} = \mu + \tau S_t + \Gamma_t$

ADF Cointegration Test[b]: $\Delta\Gamma_t = \psi\Gamma_{t-1} + \sum_{i=1}^{N} \omega_i \Delta\Gamma_{t-i} + e_t$

Coefficient Estimates[c]	*Crude Oil*	*Heating Oil*	*Unleaded Gasoline*
μ	0.1111[d]	0.0135[d]	0.0122[d]
	(3.5865)	(5.1132)	(7.1394)
τ	0.9887[d]	0.9580[d]	0.9803[d]
	(717.36)	(228.99)	(346.83)
ψ	−0.3350[d]	−0.0862[d]	−0.0897[d]
	(−9.4287)	(−11.9479)	(−5.6293)
Test:[e] $\tau = 1$ (F-statistic)	(63.55)[d]	(103.26)[d]	(49.35)[d]

[a]See notes to Table I regarding estimation periods and adjustments to futures prices.
[b]Lag length N is chosen such that Q-statistic at 24 lags indicates minimum autocorrelation in the residuals. The ADF test rejects the null hypothesis of non-cointegration if the t-statistic on the coefficient ψ is negative and below the critical values reported in Engle and Yoo (1987).
[c]Figures in parentheses below coefficient estimates are t-statistics.
[d]Indicates significance at the 1% level.
[e]To test if the slope parameter in the cointegrating regression equals unity, the cointegrating regression is estimated with an additional explanatory variable (ΔS_t), and then a standard F-test is performed [see Phillips (1991) for a description of the methodology].

that these results are based on daily closing prices and that it may not capture the intraday behavior of each market. It is likely that futures would display additional volatility in an intraday analysis. Unfortunately, intraday data for spot energy products do not exist.

Price Discovery Examination

An examination of relative price leadership between the spot and futures markets of each product is first examined in the Error Correction Model causality framework of Granger (1988). In all cases, the ECM's are estimated using Newey and West (1987) consistent covariance matrices because White (1980) tests and the Q-statistics indicate the presence of heteroskedasticity and autocorrelation (respectively) in the residuals. Table IV shows universal support for futures price leadership. For each product, the joint significance test for all relevant parameters (β and γ) is highly significant for futures explaining ΔS_t, but insignificant in spot explaining $\Delta F_t^{\prime}$. In addition, the β parameter, which represents the response of the market to errors

TABLE IV

Tests of Causality Using Error Correction Models[a]

$$\Delta S_t = \alpha + \beta\Gamma_{t-1} + \sum_{i=1}^{N} \gamma_i \Delta F_{t-i}^{\prime} + \sum_{i=1}^{N} \phi_i \Delta S_{t-i} + e_t$$

$$\Delta F_t^{\prime} = \alpha + \beta\Gamma_{t-1} + \sum_{j=1}^{M} \gamma_j \Delta S_{t-j} + \sum_{j=1}^{M} \phi_j \Delta F_{t-j}^{\prime} + e_t$$

Dependent Variable	Lags of ΔF_t, ΔS_t in ECM[b]	Estimate of Coefficient[c] β	Significance Level of Test[d]	
			All $\gamma_i = 0$	$\beta = 0$ and all $\gamma_i = 0$
Crude Oil				
ΔS_t	8	0.2999[e] (3.3560)	0.0060	0.0000
$\Delta F_t^{\prime}$	3	−0.1155[f] (2.2603)	0.7419	0.2760
Heating Oil				
ΔS_t	1	0.0858[f] (2.2851)	0.0177	0.0005
$\Delta F_t^{\prime}$	1	−0.0050 (0.2055)	0.3804	0.6801
Unleaded Gasoline				
ΔS_t	1	0.0759[f] (2.1872)	0.0437	0.0008
$\Delta F_t^{\prime}$	1	−0.0263 (1.0047)	0.4505	0.4390

[a]See Granger (1988) for description of methodology. Γ_t is the residual from the cointegrating regression of $F_t^{\prime}$ on S_t. Estimation period is Jan. 1, 1984–May 15, 1991 (Jan. 1, 1985–May 15, 1991 for unleaded gasoline). ECM's are estimated with Newey and West (1987) heteroskedasticity and autocorrelation consistent covariance matrices.

[b]Optimum lag length determined by minimizing the Schwarz (1978) criterion. To minimize problems with data snooping, lag lengths for ΔF_t and ΔS_t within each ECM are constrained to be equal.

[c]The figures in parentheses are t-statistics.

[d]These hypotheses are examined using χ^2 tests.

[e]Indicates significance at the 1% level.

[f]Indicates significance at the 5% level.

in the cointegration regression of Table III, is significantly positive for all three products when futures is employed to explain ΔS_t. The interpretation is that when there is a deviation from the equilibrium cointegrating relationship as measured by $F_t^{\prime} - \mu - \tau S_t$, it is the spot market which adjusts to restore equilibrium, implying that the futures prices lead the spot.[16] The significant coefficient for β in

[16]These results are in direct opposition to those of Quan (1992). Surprisingly, Quan's coefficient for the cointegrating parameter, Z_{t-1}, is insignificant (see his Table V, p. 146). Since Quan's data were of a monthly time interval, results from applying monthly data to the tests in this study are reported in the following section.

TABLE V
Garbade and Silber Model Estimates[a]

$$\Delta P_t = \alpha_s Ds_t + \beta_s[Ds_t * (f_{t-1} - s_{t-1})] + \alpha_f Df_t + \beta_f[Df_t * s_{t-1} - F_{t-1})] + e_{1t}$$

Parameter[b]	*Crude Oil*	*Heating Oil*	*Unleaded Gasoline*
β_s	0.2311	0.0777	0.0889
t-statistic: $\beta_s = 0$	(4.4646)[c]	(2.8581)[c]	(2.3360)[d]
β_f	0.0820	0.0133	0.0121
t-statistic: $\beta_f = 0$	(2.3571)[d]	(0.6813)	(0.3952)
θ	0.7381[d]	0.8541[e]	0.8803
χ^2 statistic: $\theta = 0.5$	(5.7139)	(3.7100)	(2.4747)
δ	0.6868	0.9090	0.8990
χ^2 statistic: $\delta = 0$	(121.2147)[c]	(728.2668)[c]	(338.6455)[c]
χ^2 statistic: $\delta = 1$	(25.2029)[c]	(7.3981)[c]	(4.2780)[d]

[a] $s_t = \ln(S_t)$; $f_t = \ln(F_t^l)$. "Stacked" spot and futures data are used. Ds_t, Df_t are 0, 1 dummy variables denoting spot and futures prices, respectively. ΔP_t reflects ΔS_t when $Ds_t = 1$ and ΔF_t when $Df_t = 1$. Figures in parentheses are t and χ^2 statistics calculated from Newey and West (1987) autocorrelation and heteroskedasticity-consistent covariance matrices. Number of observations in the stacked regression is 3522 (3046 for unleaded).
[b] $\theta = \beta_s/(\beta_s + \beta_f)$. Values greater than 0.5 indicate futures price leadership predominates.
$\delta = 1 - \beta_s - \beta_f$. A positive value less than one indicates the percentage of disequilibrium mispricing that remains in the subsequent time period.
[c] Indicates significance at the 1% level.
[d] Indicates significance at the 5% level.
[e] Indicates significance at the 10% level.

the crude oil regression on ΔF_t^l suggests that there is some feedback leadership by the spot market. However, the coefficient, −0.1165, is less than half of that showing futures leadership. In addition, the joint test of the significance of all parameters is insignificant. Consequently, futures are shown to dominate in price discovery in all three energy markets.

Further investigation on price leadership using the Garbade and Silber model is reported in Table V. The price leadership statistic, $\theta = \beta_s/(\beta_s + \beta_f)$, will be greater than 0.5 whenever futures dominates in price leadership and vice versa. As shown, this is the case for all three energy contracts, although θ is not significantly greater than 0.5 for unleaded gasoline. Identical conclusions are obtained whether analysis is performed in levels or in their natural logs. Only the natural log results are reported herein to allow comparison to other research using the GS model. Both heteroskedasticity and autocorrelation are controlled for via the Newey and West (1987) estimation technique, and by estimating the GS model in a single equation framework.[17]

The levels of θ are large; 0.7381, 0.8541, and 0.8803, for crude, heating oil, and unleaded gasoline, respectively. These levels suggest that the futures market in each commodity is the dominant market for price leadership. These values are comparable to those found in other studies that use daily data and actual spot prices. For example, Garbade and Silber report (1983, Table 2) θ levels of 0.85, 0.76, 0.54, 0.75, and 0.54 for wheat, corn, oats, orange juice, and copper markets. Oellerman et al. report (1989, Table III) θ levels of 1.00 and 0.76 for feeder cattle for two contiguous 3-year periods. Schwarz and Laatsch report (1991, Table III) evolving θ levels that are positively related to the growth in futures trading for stock index futures; $\theta = 0.1055, 0.000, 1.000$, and 0.9639 over four contiguous periods. Khoury and Yourougou (1991, Table 1) results show θ levels (recalculations of the β_c and β_f reported) of 0.953, 0.802, and 0.970 for barley, oats, and canola markets. Finally, Schroeder and Goodwin report (1991, Table I) θ levels ranging from 0.24 to 1.00 for various years of the live hog market, with 10 of 15 years showing θ levels greater than 0.50.

The estimates of δ herein suggest that a significant portion of the mispricing series extends beyond a day. For crude oil futures, the estimated δ of 0.6868 indicates that 68.68% of the mispricing lasts until the next day. For heating oil and unleaded gasoline, δ is estimated to be 0.9090 and 0.8990, respectively. For all three products, δ is significantly greater than zero, and significantly less than one, at the 1% level. However, the higher values for heating oil and unleaded gasoline suggest that the supply of arbitrage services is less in these markets than for crude oil.[18]

[17]In addition to the estimates reported in Table V, a control for seasonal price movements is implemented with a version of eq. (10) in which DF_t and DS_t are each replaced by 12 monthly dummy variables. A joint test of the equality of these coefficients fails to reject the null hypothesis (equality) for crude oil and unleaded, indicating that seasonality is not a significant factor for these products. Substantial seasonality for heating oil is found, but the estimated β_s and β_f coefficients, and their associated standard errors do not change much when the seasonal dummies are included in the regression.

[18]Estimates of δ for other futures contracts on daily data include: (1) Garbade and Silber (1983); wheat, 0.97; corn, 0.96; oats, 0.96; orange juice, 0.84; and copper, 0.92; (2) Oellerman et al. (1989), feeder cattle, 0.95 and 0.97 for 1979–1982 and 1983–1986, respectively, (3) Schwarz and Laatsch (1991), MMI equity futures, 0.51, 0.61, 0.00, and 0.09 for 9/85–6/86, 7/86–4/87, 5/87–9/87, and 11/87–3/88, respectively, (4) Khoury and Yourougou (1991), barley, 0.92; oats, 0.89; and canola, 0.97, and (5) Schroeder and Goodwin (1991); live hogs, ranging from 0.72 to 1.00 for individual years from 1975 to 1989. The comparison with the results herein suggests that crude oil futures are arbitraged relatively well. The degree of arbitrage in the two refined product markets is similar to levels found for agricultural commodities.

Weekly and Monthly Results

Both weekly and monthly price series are constructed from the daily series to examine the impact of time upon arbitrage and price leadership.[19] A priori, the model of Garbade and Silber suggests that a longer time interval would mobilize a greater amount of arbitrage capital to profit from market mispricings. Given that the levels of δ for daily observations are significantly greater than zero, additional arbitrage capital does indeed materialize so that estimates of δ are lower with longer measurement intervals.

The daily δ results for crude oil imply that only 15.28% (0.6868^5) and 0.03% (0.6868^{22}) of the mispricing remains after 1 week (5 trading days) and after 1 month (22 trading days). This suggests that much of the price leadership in these markets is of an interval *less* than one month. Results in Table VI support this supposition. For crude oil, the δ statistic falls first to -0.0251 for weekly observations and then reverses and rises to 0.4891 for monthly observations. This instability is likely a result of too long of a time interval between price measurements. If price leadership is of a shorter duration, then a long time interval of measure will add noise to the estimation results. Confirmation of this supposition can be seen by the increasing number of non-meaningful figures that result with monthly data. The GS model theoretically limits both β_s and β_f to be bounded between 0 and 1. However, for the weekly and monthly data, many of the estimated parameters exceed 1 or are below 0. There are no violations using daily data.

In comparing these monthly results with Quan (1992), some differences remain. Quan reports (Table VI, Case 1) β_s (β_p in his article) $= 0.003$, $\beta_f = 0.480$, $\theta = 0.006$, and $\delta = 0.52$. These are contrary to the results reported in Table VI of this study where β_s is significantly positive and β_f is insignificantly different from zero. Rerunning the analysis using the period of Quan's study individually for both F_t and F_t', and using OLS and Newey and West (1987) techniques does not change the conclusions herein.

[19]The weekly dataset is created using Tuesday quotations, or the preceding day if Tuesday is a holiday. The number of observations is 385 for crude and heating oil, and 332 for unleaded gasoline. Weekly regressions are estimated using Newey and West (1987) consistent covariance matrices. The monthly dataset is constructed using prices for the first deferred contract on the day when the nearby contract ceases trading. Thus, the monthly price series represents futures contracts with approximately one month before expiration. The use of mid-month observations as in Quan (1992) does not yield materially different results from those reported in this study.

TABLE VI
Garbade and Silber Model With Weekly and Monthly Data[a]

Parameter[b]	*Crude Oil*	*Heating Oil*	*Unleaded Gasoline*
Panel A: Weekly Data[b]			
β_s	0.6208	0.3532	0.2423
t-statistic: $\beta_s = 0$	(3.6266)[c]	(3.0980)[c]	(1.9728)[c]
β_f	0.4043	−0.1431	0.1246
t-statistic: $\beta_f = 0$	(2.4174)[d]	(1.4387)	(1.1744)
θ	0.6056	1.6816[e]	0.6604
χ^2 statistic: $\theta = 0.5$	(0.8178)	(10.7533)[c]	(0.5262)
δ	−0.0251	0.7900	0.6331
χ^2 statistic: $\delta = 0$	(0.0110)	(27.2489)[c]	(5.2257)[c]
χ^2 statistic: $\delta = 1$	(18.3469)[c]	(1.9260)	(5.1140)[d]
Panel B: Monthy Data[f]			
β_s	0.8579	1.1398	1.1229
t-statistic: $\beta_s = 0$	(2.4519)[d]	(6.2191)[c]	(2.9059)[c]
β_f	−0.3489	−0.5397	−0.6646
t-statistic: $\beta_f = 0$	(1.0709)	(3.3916)[c]	(1.8707)[g]
θ	1.6829[e]	1.8990[e]	2.4501[e]
F-statistic: $\theta = 0.5$	(6.3764)[d]	(47.8823)[c]	(11.5956)[c]
δ	0.4891	0.3998	0.5418
F-statistic: $\delta = 0$	(1.0445)	(2.7137)	(1.0652)
F-statistic: $\delta = 1$	(1.1393)	(6.1146)[d]	(0.7620)

[a]See Table V for description of model estimated. Figures in parentheses are t, χ^2, and F statistics as indicated in column 1.
[b]Weekly dataset contains Tuesday price quotations for spot and futures (preceding trading day used if Tuesday is a holiday). Number of observations: 385 for crude oil and heating oil, 332 for unleaded gasoline; the stacked single equation procedure doubles these numbers for estimation purposes. Weekly regressions are estimated using Newey and West (1987) consistent covariance matrices.
[c]Indicates significance at the 1% level.
[d]Indicates significance at the 5% level.
[e]Garbade and Silber theoretical procedure requires both β_s and β_f to be between 0 and 1. θ is difficult to interpret when the estimated β_f is negative, or when β_s is greater than one since these conditions violate the assumptions of the GS model.
[f]Monthly dataset are constructed by sampling daily prices for the first deferred contract on the day when the nearby contract ceases trading; thus, in the dataset, the futures price is for a contract with approximately one month left until expiration. Number of observations: 87 for crude and heating oil, 75 for unleaded. Monthly regressions are estimated via OLS because neither heteroskedasticity nor autocorrelation is found in the residuals at this frequency.
[g]Indicates significance at the 10% level.

SUMMARY AND CONCLUSIONS

The temporal properties of three energy products and their futures markets—light sweet crude oil, heating oil #2, and unleaded gasoline—are examined and are found to be nonstationary with unit roots. Furthermore, the spot and futures of each product are found to be cointegrated as would be expected from the relation between cointegration and arbitrage activity.

Contrary to the earlier results of Quan (1992), the findings from error-correction [Granger (1988)] and Garbade and Silber (1983) models suggest that light sweet crude oil futures dominate in price discovery relative to its deliverable spot instrument. This analysis is based on data of higher frequency than is used by Quan. Also, the data used are the most accurate measure of spot data available to the industry itself. It is suspected that both the choice of data and the time interval of analysis are important factors underlying Quan's conclusions that "...the crude oil futures market does not play an important role in price discovery."

This study strongly suggests that futures dominate in price discovery in all three petroleum product markets. Results concerning the supply of arbitrage services in these markets are similar to those of other commodities where the price discovery role of futures has been fully established. These findings are consistent with the tremendous growth and proliferation of trading in petroleum futures.

BIBLIOGRAPHY

Baillie, R.T., and Bollerslev, T. (1989): "Common Stochastic Trends in a System of Exchange Rates," *Journal of Finance,* 44(1):167–182.

Bessler, D.A., and Covey, T. (1991): "Cointegration: Some Results on U.S. Cattle Prices," *The Journal of Futures Markets,* 11(4):461–474.

Bopp, A. E., and Lady, G. (in press): "A Comparison of Petroleum Futures Versus Spot Prices as Predictors of Prices in the Future," *Energy Economics.*

Bopp, A.E., and Sitzer, S. (1987): "Are Petroleum Futures Prices Good Predictors of Cash Value," *The Journal of Futures Markets,* 7:705–719.

Cho, D.W., and McDougall, G.S. (1990): "The Supply of Storage in Energy Futures Markets,: *The Journal of Futures Markets,"* 10(6):611–621.

Engle, R.F., and Granger, C.W.J. (1987): "Co-integration and Error Correction, Representation Estimation and Testing," *Econometrica,* 55:251–276.

Engle, R.F., and Yoo, B.S. (1987): "Forecasting and Testing in Co-Integrated Systems," *Journal of Econometrics,* 35(1):143–161.

Fuller, W. (1976): *Introduction to Statistical Time Series,* New York: Wiley.

Garbade, K.D., and Silber, W.L. (1983): "Price Movement and Price Discovery in Futures and Cash Markets," *Review of Economics and Statistics,* 65:289–297.

Granger, C.W.J. (1988): "Some Recent Developments in a Concept of Causality," *Journal of Econometrics,* 39(1/2):199–212.

Hall, S.G. (1989): "Maximum Likelihood Estimation of Cointegration Vectors," *Oxford Bulletin of Economics and Statistics,* 51(2):213–218.

Hirschfeld, D.J. (1983): "A Fundamental Overview of the Energy Futures Market," *The Journal of Futures Markets,* 1:75–100.

Johansen, S. (1988): "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control,* 12:231–254.

Johansen, S. (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica,* 59(6):1551–1580.

Khoury, N.T., and Yourougou, P. (1991): "The Informational Content of the Basis: Evidence from Canadian Barley, Oats, and Canola Futures Markets," *The Journal of Futures Markets,* 11(1):69–80.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992): "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics,* 54:159–178.

Ma, C.K., Mercer, J.F., and Walker, M.A. (1992): "Rolling Over Futures Contracts: A Note," *The Journal of Futures Markets,* 12(2):203–218.

Newey, W.K., and West, K.D. (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* 55(3):703–708.

Ollermann, C.M., Brorsen, B.W., and Farris, P.L. (1989): "Price Discovery for Feeder Cattle," *The Journal of Futures Markets,* 9(2):113–121.

Quan, J. (1992): "Two Step Testing Procedure for Price Discovery Role of Futures Prices," *The Journal of Futures Markets,* 12(2):139–149.

Phillips, P.C.B. (1991): "Optimal Inference in Cointegrated Systems," *Econometrica,* 59(2):283–306.

Schroeder, T.C., and Goodwin, B.K. (1991): "Price Discovery and Cointegration for Live Hogs," *The Journal of Futures Markets,* 11(6):685–696.

Schwarz, G. (1978): "Estimating the Dimension of a Model," *Annals of Statistics,* 6:461–464.

Schwarz, T.V., and Laatsch, F.E. (1991): "Dynamic Efficiency and Price Leadership in Stock Index Cash and Futures Markets," *The Journal of Futures Markets,* 11(6):669–683.

Sephton, P.S., and Larsen, H.K. (1991): "Tests of Market Efficiency: Fragile Evidence from Cointegration Tests," *Journal of International Money and Finance,* 10:561–570.

Serletis, A., and Banack, D. (1990): "Market Efficiency and Cointegration: An Application to Petroleum Markets," *Review of Futures Markets,* 9(2):372–385.

Stoll, H.R., and Whaley, R.E. (1990): "The Dynamics of Stock Index and Stock Index Futures Returns," *Journal of Financial and Quantitative Analysis,* 25(4):441–468.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica,* 48:817–838.