Price Dynamics and Error Correction in Stock Index and Stock Index Futures Markets: A Cointegration Approach

Mahmoud Wahab
Malek Lashgari

INTRODUCTION

The temporal relation between stock index spot and futures markets has been, and continues to be, of interest to academicians, regulators, and practitioners alike for a variety of reasons. First, the issue is inextricably linked to few central notions in financial theory, notably market efficiency and arbitrage. In perfectly efficient markets, profitable arbitrage should not exist as prices adjust instantaneously and fully to new incoming information. Therefore, new information disseminating into the marketplace should be immediately reflected in spot and futures prices by triggering trading activity in one or both markets simultaneously so that there should be no systematic lagged responses long enough, or large enough to economically exploit, considering transaction costs. Second, it is often believed that futures markets potentially provide an important function of price discovery. If so, then futures prices or movements thereof should contain useful information about subsequent spot prices; beyond that already embedded in the current spot price. A third issue pertains to potential volatility-spillover effects of futures trading. Specifically, these markets have long been suspected of exerting a destabilizing influence on the underlying spot market. Although still largely unsettled at both the theoretical and empirical levels, there is some evidence that trade in index futures increases cash market volatility [see, for example, Harris (1989)]. Essentially, proliferation of index arbitrage activities may have further exacerbated stock market volatility. However, it can also be argued that index arbitrage activities through program trading may have served to make the stock market more efficient, causing prices to adjust more rapidly to new information, and help keep stock index and stock index futures price movements more synchronous. Indeed, this observation is consistent with the reduced transaction costs,

We gratefully acknowledge extremely helpful comments from two anonymous referees. Any remaining errors are our responsibility.

Mahmoud Wahab is an Assistant Professor of Finance at the University of Hartford.

Malek Lashgari is an Associate Professor of Finance at the University of Hartford.

The Journal of Futures Markets, Vol. 13, No. 7, 711–742 (1993)

CCC 0270-7314/93/070711-32

increased financial leverage opportunities, and absence of short-sale constraints in futures markets, which make trading in stock index futures more attractive than trading in the market for the underlying stocks.

As is widely known, index arbitrage is motivated by deviations of futures prices from their perceived fair value estimates as established perhaps by the familiar cost of carry model. These deviations may persist due to imperfect substitutability between spot and futures markets; the differential speed with which spot and futures markets reflect new information; divergent opinions of market participants regarding parameter values of the valuation model, transaction costs, regulatory constraints, and other market imperfections. Potentially, the differential response of both markets gives rise to a lead–lag structure which may be influenced by several factors, some of which are spurious, with no economic significance; and some are institutional, reflecting the different characteristics of both markets. On the one hand, changes in index futures prices may spuriously exhibit an empirical lead over spot index price changes because of infrequent trading of stocks within the index. Essentially, component stocks may not trade in every instant so that observed prices may not reflect "true" prices. In turn, the spot index would not update actual developments in the component stocks, thereby lagging actual developments in the stock market. An index futures contract, however, represents a single claim, as opposed to a portfolio of component securities and, therefore, should not suffer from the asynchronous trading problem observed for the spot index [Stoll and Whaley, (1990)].

On the other hand, other characteristics of both markets may also contribute to an observed lead of index futures over the spot index. For example, time delays in updating and transmitting updated spot index values, in addition to differential transaction costs which favor futures-market transacting over trading in the spot market may explain the tendency of stock index futures price changes to lead spot index price changes. Stoll and Whaley (1990), describe three possible time delays: (i) delays in entering stock index transactions into the computer, (ii) delays in updating and communicating new spot index values, and (iii) delays in recording the updated spot index value at the futures exchange. Since long and short positions can be established more easily, and less expensively in futures markets, more so than in the spot market, trading based on revised expectations can be undertaken more frequently, and with less cost in the futures markets. Therefore, futures prices may move first, followed by spot price movements in response to changes in consensus expectations about the stock market. A stronger lead from spot-to-futures may not, however, be inconceivable since the value of the spot index and its more recent changes represent part of the information set used by futures traders. Changes in the spot index (or cumulative changes thereof) may induce changes in the futures market sentiment that would be reflected in subsequent futures price changes, giving rise to a tendency of the index futures to lag the spot index.

This study reexamines empirically the daily price-change relation between stock index and stock index futures markets for the Standard and Poor 500 index (S&P 500) and the Financial Times–Stock Exchange 100 share index of London (FT–SE 100). Recent developments in the theory of cointegration are employed to test the causal relationships between time series. In particular, Engle and Granger (1987) showed that if two series are each nonstationary, but that a linear combination of the two variables is stationary so that both are cointegrated, then a bivariate dynamic model that uses only first differences (and lagged changes thereof) is misspecified because it ignores interim short-run adjustments to long-run equilibrium.[1] Additionally, a percent-change

[1]Essentially, taking first differences eliminates all long-run information in the data which, in turn, is recovered by including an equilibrium error term.

specification would be misspecified also because it fails to consider short-run dynamics by leaving out lagged values. On the other hand, a level specification is simply a long-run equilibrium relation that does not account for the short-run dynamics in and between the series. Lead and lag results reported in earlier studies, therefore, may be simply an artifact of a particular model specification. Thus, this article extends the current studies of lead and lag relationships between stock index and stock index futures prices by applying cointegration analysis to investigate the robustness of the previously reported daily evidence to an alternative model parameterization. The methodology employed considers non-stationarity, and allows for both short-term and long-term adjustments.

The link between cointegration and causality stems from the fact that if spot and futures indexes are cointegrated, then causality must exist in at least one direction, and possibly in both directions [Granger, (1986)]. Since cointegration implies that each series can be represented by an error correction model that includes last period's equilibrium error as well as lagged values of the first differences of each variable, temporal causality can be assessed by examining the statistical significance, and relative magnitudes, of the error correction coefficients and the coefficients on the lagged variables.

Significant causal relationships would, however, be incompatible with market efficiency because they would imply that forecast accuracy of the spot (futures) market's subsequent performance can be improved upon by using past information from the futures (spot) market. To avoid contradicting the efficient markets paradigm, the joint comovement of price changes in the two markets should be predominantly contemporaneous. Combining the above two notions would require that two prices from a pair of efficient markets cannot be cointegrated. Therefore, whether or not the S&P 500 spot and futures indexes are cointegrated is tested in this study. To investigate the robustness of the phenomenon of cointegration of a pair of cash and futures stock indexes outside the U.S., the joint behavior of the U.K. FT–SE 100 spot and futures indexes is similarly analyzed. Appropriately specified error correction models are then estimated if cointegration is "accepted." Such models allow the long-run components of variables to obey equilibrium constraints, while short-run components have a flexible dynamic specification which may diverge from the long-run trend. They also indicate not only the proportion of disequilibrium from one period that is corrected in the next period, but also, the relative magnitude of adjustments in both markets towards equilibrium. This also exposes the nature of causal relationships.

The plan of this article is as follows: It briefly describes the general relation between stock index spot and futures prices and contains a review of the more recent literature on the temporal relation between stock index and stock index futures markets. It then defines cointegration, its relationship to causality, and what it implies for market efficiency; describes the data and methodology; and presents the empirical results.

THE RELATIONSHIP BETWEEN STOCK INDEX AND STOCK INDEX FUTURES PRICES

The continuous time representation of the theoretical fair value estimate of the stock index futures (forward) price is fairly approximated by the familiar net cost-of-carry model: $F_t = S_t e^{(r-d)(T-t)}$ where F_t is the index futures price at time t; S_t is the spot index price at t; r is the continuously compounded cost of carrying the spot index basket from the present t, to time T which is the expiration date of the stock index futures contract. $T - t$ is the time remaining to expiration of the futures contract and d is the dividend yield on the stock index. $r - d$ represents the net cost of carry which is the time value cost of wealth tied up in the stock index investment, offset by the flow of dividends from the index. The linkage between prices of the underlying basket of stocks and stock index futures are thought of as being maintained by professional arbitrageurs who employ computerized trading systems to capitalize on deviations of stock index futures prices from "perceived" fair values, or alternatively, whenever violations of the above parity relation arise. If actual futures prices exceed their perceived fair value, the futures contract is overvalued. This justifies a long arbitrage position in which the futures contract is sold and the stock index portfolio (or a close replica thereof) is bought. This is financed with riskless borrowing which allows an arbitrageur to lock-in an arbitrage profit of $F_t - S_t e^{(r-d)(T-t)}$. On the other hand, an undervalued futures contract triggers a short arbitrage position where an arbitrage profit of $S_t e^{(r-d)(T-t)} - F_t$ can be earned by buying futures, selling the underlying basket of stocks, and investing the sale proceeds at the riskless rate. One empirical implication of the cost-of-carry model is that only contemporaneous values of the parameters enter the model, so that, in perfectly efficient and continuous spot and futures markets, price adjustments are instantaneous. In other words, assuming that the cost-of-carry model is a true characterization of the spot–futures price relationship and that markets are perfect and efficient, exclusion restrictions on lead and lagged values of spot and futures prices should hold at stringent significance levels when using an empirical form of the model. The observed relation between price changes in the two markets will be noisy due to market imperfections and because price observations from the two markets are not simultaneous. Therefore, the normal relationship between stock index and stock index futures prices would be bounded from above and below by a no-arbitrage trading band whose width is determined by the consensus perception of market imperfections. Only when the spot–futures price differential moves outside the no-arbitrage boundaries will program purchase and sales be triggered to lock-in an arbitrage profit. However, it would be very difficult to a priori establish the width of the no-arbitrage trading band because it is dictated by a divergence of opinions among market participants regarding the valuation parameters. Furthermore, it may be time-varying [Mackinlay and Ramaswamy, (1988)]. In addition, if there are economic incentives for traders to use one market over the other, a lead–lag relation between price changes in the two markets is likely.

LITERATURE REVIEW

Earlier studies that examine lead–lag relationships between stock index spot and futures prices used Granger's (1969) notion of causality which rests on an incremental predictability criterion. A variable, X, causes another variable, Y, if the past history of X can be used to predict Y more accurately than simply using the past history of Y alone. At the same time, the past history of Y cannot help predict X more accurately, once the history of X and other relevant information are accounted for. Several test procedures have been employed to investigate the lead–lag relation. The most popular are: a one-sided distributed-lag model [Granger (1969)],[2] a two-sided distributed-lag model [Sims

[2]Granger's (1969) one-sided regression approach involves estimating the following bivariate autoregressive model (after appropriate pre-filtering), using empirically determined filters:

$$Y_t = \sum_{k=1}^{n} \alpha_k Y_{t-k} + \sum_{i=1}^{m} \beta_i X_{t-i} + \mu_t$$

$$X_t = \sum_{k=1}^{n} \gamma_k Y_{t-k} + \sum_{i=1}^{m} \delta_i Y_{t-i} + \mu_t'$$

Unidirectional causality from X to Y requires that some of the β_i must be nonzero, while all δ_i must be equal to zero. Similarly, for Y to cause X, some of δ_i must be nonzero while all β_i must be jointly zero.

(1972)],[3] and a cross-correlation technique [Haugh (1976)].[4] Evidence that the futures market tends to lead the stock market has been reported in a number of studies. Kwaller, Koch, and Koch (1987) examined the intra-day price relationship between S&P 500 futures and the S&P 500 index. Their results show that both S&P 500 spot and futures prices are simultaneously related on a minute-to-minute basis throughout the trading day, but that a lead–lag relation also exists. The lead from futures-to-cash however, appears to be more pronounced relative to the lead from cash-to-futures markets. Their test is in the spirit of Granger's one-sided regression approach, with instantaneous interactions explicitly considered.

Stoll and Whaley (1990) investigated causal relationships between spot and futures prices using intra day data for both S&P 500 and the Major Market Index (MMI). To control for nonsynchroneity and bid–ask bounce effects, they prefiltered spot index returns using an ARIMA filter. Stock index return innovations were then used in a two-sided regression procedure, using lead, contemporaneous, and lagged returns on the nearby futures contract. Feedback was detected, but the futures lead was stronger than the spot index lead.

Ng (1987) used the cross-correlation function approach to test for causal relationships between spot and futures prices for the S&P 500, the Value Line Index (VLI), and five exchange rates. Employing daily data for about 5 years, Ng reported evidence that futures prices lead rather than lag spot prices by one day, although the lead coefficients were

[3]The two-sided regression approach of Sims (1972) involves estimating a two-sided regression model of the form:

$$X_t = \sum_{i=-m}^{+m} \beta_i Y_{t-i} + \omega_t$$

and test the hypothesis that all coefficients on future Y values are zero. An analogous regression of Y_t on past and future X_t is estimated to test if Y causes X. Since F-tests on estimated coefficients are invalid when errors are serially correlated, the X and Y series are usually prefiltered before estimating and testing the model. The F-test is employed to test exclusion restrictions on the appropriate parameters. For alternative tests of linear dependence and feedback between time series, see for example, Gweke (1982), and Gweke, Meese, and Dent (1983).

[4]The Haugh (1976) approach to testing causality involves two steps. First, Box–Jenkins univariate ARIMA models are used to estimate appropriate filters for each series and to compute in-sample innovations. Second, estimated innovations are used to compute the sample cross-correlation function to make inferences about the population cross-correlation function which, in turn, permits inferences about causal relationships in the system. The population cross-correlation function is constructed for different leads and lags k, where $k = -m, +m$, and is given by:

$$\rho_{\mu v}(k) = \gamma_{\mu v}(k)/\left[\gamma_v^2(0) \cdot \gamma_\mu^2(0)\right]^{1/2}$$

where $\gamma_{\mu v}(k) = E(\mu_{t-k} v_t)$, $\gamma_\mu^2 = E(\mu_t^2)$, and $\gamma_v^2 = E(v_t^2)$. Since the true cross-correlations are of course unknown, sample cross-correlations are used. Under the null of independence (no causality), the sample cross-correlation function has zero values at all positive and negative lags. The large-sample test statistic (n relative to the lag order k) is given by:

$$S = n \sum_{k=-m}^{+m} (r_{\mu v}(k))^2$$

while the small-sample test statistic is:

$$S = n^2 \sum_{k=-m}^{+m} (n - |k|)^{-1} (r_{\mu v}(k))^2$$

where $\gamma_{v\mu}(k)$ denotes the estimated cross-correlation coefficient at lag (k), n is the number of observations, and k is the lag order. The test statistic is distributed as χ^2 variate with $(2m + 1)$ degrees of freedom; with symmetric lags employed on either side.

rather weak in magnitude. No lead was detected for spot prices. Ng used Haugh's test, but not the "modified" Haugh's test [Koch and Yang, (1986)], that accounts for possible serial correlation in the cross-correlation function.[5]

More recently, Chan, Chan, and Karolyi (1991) examined, simultaneously, the interdependence in price change and price-change volatility. They found much stronger bidirectional dependence between stock index and stock index futures price changes when the volatility of price changes is also considered. Their evidence is consistent with the hypothesis that both markets serve important price discovery roles. They used a bivariate AR(1)–GARCH(1,3) model of the joint process, which represents an extension to the autoregressive conditional heteroskedastic (ARCH) family of models developed by Engle (1982) and generalized (GARCH) by Bollerslev (1986).

COINTEGRATION AND CAUSALITY

Cointegration theory implies that for a vector of time series, the variables are said to be cointegrated if linear combinations thereof are stationary without differencing, even though the vector elements need to be differenced at least once to become individually stationary. If a series needs to be differenced d times before it becomes stationary, it is said to be integrated of order d, denoted as $X_t \sim I(d)$. If two time series X_t and Y_t are each individually $I(d)$, then in general, most linear combinations of X_t and Y_t such as $z_t = X_t - aY_t$ are also nonstationary. If differencing d times causes X_t and Y_t to become stationary, then z_t will also be stationary. If a linear combination of X_t and Y_t exists that is stationary, then the two series X_t and Y_t are said to be cointegrated of order zero. Therefore, in the cointegrating or equilibrium regression $X_t = aY_t + z_t$, (a) is the cointegrating parameter; or alternatively, for a vector of time series, (**a**) is called the cointegrating vector. If two series are covariance stationary without trends-in-mean,[6] and are cointegrated, then there exists an error correction represenction for each series which is not liable to the problems of spurious regression [Granger and Newbold, (1974)].

If ΔX_t and ΔY_t denote, respectively, first differences in spot and futures prices one time; then reverse the designation so that spot and futures first differences are now denoted by ΔY_t and ΔX_t, respectively, the following four error correction specifications

[5]The Haugh test may be weak if X and Y are related over long distributed lags but in which individual lag coefficients are small. More specifically, the Haugh test would be unable to distinguish one cross-correlation function between two variables that are independent from another cross-correlation function where variables are related over a long distributed lag with estimated coefficients that are small in magnitude but are arranged in a distinct pattern. Thus, Haugh's test ignores autocorrelation in successive cross-correlation coefficients [Koch and Yang (1986)]. A more powerful test statistic is given by:

$$r_i^* = n \sum_{k=-m}^{+m} \left[\sum_{p=0}^{i} r_{\mu v}(k + p) \right]^2$$

where m denotes the lead–lag length used in estimating the cross-correlation function, p is the autocorrelation lag order in the cross-correlation function, and other variables are as defined earlier. This test statistic can be easily shown to incorporate the Haugh statistic as a special case when $p = 0$.

[6]In Engle and Granger (1987), it is assumed that all deterministic components are removed before the analysis. Engle and Yoo (1987) showed that even if deterministic trends are not removed for each series beforehand, the same linear combination which eliminated nonstationarity, will also eliminate the linear trend.

and two cointegrating regressions are possible:[7]

$$\Delta S_t = \alpha_0 - a_1 z_{t-1} + \sum_{i=1}^{m} c_{i,s}\Delta S_{t-i} + \sum_{j=1}^{m'} d_{j,f}\Delta F_{t-j} + e_{s,t} \tag{1}$$

$$\Delta F_t = \alpha_0' - a_2 z_{t-1} + \sum_{i=1}^{n} c_{i,s}'\Delta S_{t-i} + \sum_{j=1}^{n'} d_{j,f}'\Delta F_{t-j} + e_{f,t} \tag{1a}$$

where

$$z_t = S_t - [b + aF_t] \tag{1b}$$

and

$$\Delta F_t = \alpha_1 - a_3 z_{t-1}' + \sum_{i=1}^{p} \lambda_i \Delta F_{t-i} + \sum_{j=1}^{q} \delta_j \Delta S_{t-j} + e_{f,t}' \tag{2}$$

$$\Delta S_t = \alpha_1' - a_4 z_{t-1}' + \sum_{i=1}^{p'} \lambda_i' \Delta F_{t-i} + \sum_{j=1}^{q'} \delta_j' \Delta S_{t-j} + e_{s,t}' \tag{2a}$$

where

$$z_t' = F_t - [b' + a'S_t] \tag{2b}$$

and where $e_{s,t}$, $e_{f,t}$, $e_{f,t}'$, and $e_{s,t}'$ are white noise processes, possibly contemporaneously correlated, and with, for example, $|a_1| + |a_2| \neq 0$. It is noted that both variables are treated as jointly endogenous, so that there is no a priori choice of which variable corresponds to X_t and which to Y_t. Furthermore, the lag orders are not necessarily constrained to be the same across the two equations. Therefore, for two variables, there are four possible error correction specifications. For example, eq. (1) has the interpretation that the change in S_t is due to both "short-run" effects, possibly from both ΔFs and ΔSs, and to last-period equilibrium error (from the cointegrating regression) which represents adjustment to long-run equilibrium. The coefficient (a) attached to the error-correction term measures the single-period response of the LHS variable to departures from equilibrium. If this coefficient is small, then the LHS variable (ΔS_t) has little tendency to adjust to correct a disequilibrium situation, so that most of the adjustments may be accomplished by ΔF_t in eq. (1a). The error correction terms (z_t) and (z_t') enter into the four equations with a one-period lag and are estimated from the cointegrating regressions, with constant terms being included to make the mean of the error series, zero. Equations (1b) and (2b) may be called the forward and reverse cointegrating regressions, respectively.

In terms of eqs. (1) and (1a), unidirectional causality from spot-to-futures requires: (i) that some of the $c_{i,s}'$'s must be non-zero while all the $d_{j,f}$'s must be equal to zero, and/or (ii) that the error correction coefficient a_2 in eq. (1a) is statistically significant at conventional levels.[8] Similarly, in terms of eqs. (1) and (1a), for futures market price

[7]The need to reverse the X, Y designations stems from the nonuniqueness of the parameter estimates. Essentially, a least squares fit of a reverse cointegrating regression will not simply give the reciprocal of the coefficient in a forward cointegrating regression.

[8]Equations (1) and (1a) could be extended further to include the contemporaneous variables ΔY_t and ΔX_t in the two equations, respectively, so that they form a complete dynamic simultaneous equation system [see for example, Hall (1986)]. This may be particularly relevant for this investigation since the broad evidence on the lead–lag relation between spot and futures prices indicates that the two markets are simultaneously related as well. The estimation procedure in this case will have to be modified to avoid simultaneous-equation-bias.

changes to Granger-cause subsequent changes in the spot index, some $d_{j,f}$'s must be non-zero, all $c^{l}_{i,s}$'s are individually, and jointly zero, and the error correction coefficient a_2 must be zero or negligible in magnitude. A similar reasoning can be used to detect causal relations between spot and futures prices when using eqs. (2), (2a), and (2b). In these cases, the error correction series represents, in a sense, a futures index shock.

In terms of the entire set of eqs. (1) to (2b), if the error terms enter significantly into the error correction eqs. (1a) and (2a), then neither variable can be considered weakly exogenous, and feedback is said to exist, so that evidence on causal ordering of the two variables is inconclusive. In sum, the error correction coefficients serve two purposes: to identify the direction of causal relation between two time series [eqs. (1a) and (2a)] and to show the speed with which departures from equilibrium are corrected in the short run by changes in the LHS variable [eqs. (1), (1a), (2), and (2a)]. The magnitude of the error correction coefficient in each of eqs. (1), (1a), (2), and (2a) also has some implications insofar as market efficiency is concerned. Basically, in perfectly efficient and continuous spot and futures markets, disequilibrium in each market (which may indicate inefficiency) or between spot and futures prices (which may indicate possible index arbitrage opportunities) should not occur. Alternatively, any disequilibrium that might signal an opportunity for economic profits should be fully and instantaneously dissipated. Hence, with additional parameter restrictions, informational efficiency may be verified. For spot and futures markets to be considered efficient, the following joint restriction must be satisfied: (i) the error correction coefficient is equal to unity, and (ii) any lagged coefficients are jointly zero in each error correction equation.[9]

DATA AND METHODOLOGY

The data used in this study are daily closing spot and futures prices for both the Standard and Poor 500 index and the Financial Times index, for the period between January 4, 1988, and May 30, 1992. The S&P 500 spot and futures data are collected from *The Wall Street Journal,* while the FT–SE 100 spot and futures data are obtained from *The Financial Times* of London. Stock index futures prices are always those of the nearby contract. The FT–SE 100 stock index incorporates the top 100 British companies. The value of the index represents the average share price weighted by market capitalization of the largest 100 firms listed on the London Stock Exchange. The index is calculated as a weighted arithmetic index with a base value of 1000 starting on January 3, 1984 and is traded between 0900–1540 (GMT). Trading is conducted in 2-week account periods (3-week periods if a public holiday occurs). Settlement occurs on account day which is always the second Monday after the end of the account period. Accounts are normally 10 business days in duration. On the account day, or day of settlement, payment is tendered or shares are delivered. In early 1984, the London International Financial Futures Exchange (LIFFE) introduced the FT–SE 100 stock index futures contract. The contract is cash-settled. Settlement is made against the average of the FT–SE 100 spot index level between 11:10 A.M. and 11:20 A.M. on the last trading day, excluding the highest and lowest levels. Each futures index point has a value of £25, and thus, the contract's notional value is obtained by multiplying the futures-index level by £25. The delivery cycle is quarterly, with delivery months of March, June, September, and December of

[9]Because markets are not frictionless, the unit restriction on the error parameter, when ΔS_t is the dependent variable, is more than likely to be always rejected. This is especially true in the spot market due to short sale restrictions and higher transaction costs.

each year. The delivery day is the first business day after the last trading day. Stock index futures are traded between approximately 9:05 A.M. and 4:05 P.M. (GMT).

Engle and Granger (1987) suggest a two-step estimation procedure to test for cointegration. This procedure gives coefficient estimates that rapidly converge on the true parameter values so that the estimated cointegration parameter is highly consistent. In the first step, a levels-regression is performed to generate residuals which may be thought of as equilibrium pricing errors. Residuals are then subjected to a variety of tests for cointegration. If cointegration is "accepted," the second step involves entering residuals into the error correction models in place of the levels terms. This two-stage procedure has the effect of imposing the cointegrating parameter (vector) on the error correction model. In this article, several of the familiar tests for cointegration are undertaken [see Engle and Granger (1987)]. The first test, the Cointegrating Regression Durbin–Watson (CRDW) statistic, uses residuals estimated from the cointegrating regressions. If this statistic is sufficiently large, the residuals are declared stationary, and X_t and Y_t will be cointegrated. The second and third tests use the Dickey–Fuller [Dickey and Fuller (1979) and (1981)], and the Augmented Dickey–Fuller test to examine whether estimated residuals are stationary. If there is a unit root in each series, then X_t and Y_t are not cointegrated. The appropriate test statistic for the DF and ADF tests is the pseudo t-statistic whose critical value can be computed using the procedure outlined in Mackinnon (1991). The fourth test involves estimating a Restricted Vector Autoregression model (RVAR). The fifth test estimates an Augmented Restricted VAR (ARVAR). In the above tests, cointegration constraints are superimposed since the resulting test statistics are conditional on both the estimate of the cointegrating parameters (a) or (a′) from eq. (1b) or eq. (2b) and the equilibrium error generated by the cointegrating regression. Therefore, two additional models—an Unrestricted Vector Autoregression (UVAR) and an Augmented Unrestricted VAR (AUVAR)—are estimated using the levels specification. For both pairs of assets, all of the above tests are conducted twice with the designation of the X_t and Y_t variables reversed. The null hypothesis is always taken to be that of non-cointegration.

EMPIRICAL RESULTS

Time Series Stationarity

Traditional statistical tests used for inference presume the use of stationary data. Regressing non-stationary variables onto each other leads to potentially misleading inferences about the estimated parameters and the degree of association. Therefore, the first step involves testing for the presence of a unit root in the autoregressive representation of each index. Unit root tests provide an easy method of testing whether a series is stationary,[10] so that rejection of the unit root hypothesis is necessary to support stationarity.[11] More importantly, unit root tests are needed to establish whether the series are integrated with a uniform order of integration, a requirement for Engle and Granger's (1987) tests for cointegration. For each of the S&P 500 and FT–SE 100 spot and futures

[10]Examination of plots of the series and its differences, and inspection of the sample autocorrelation function (ACF) of the series and its differences for failure to damp out quickly, although useful in making decisions about the degree of differencing needed to achieve stationarity, often run into difficulties. This is true especially in borderline cases if the series is stationary. But the lag polynomial contains factor(s) close to unity $((1 - \phi B) \approx (1 - B))$ thus, making it difficult to detect non-stationary behavior. In addition, employing such methods on a stand-alone basis usually entails subjective judgement in the identification process.

[11]Since unit root tests do not test for constant moments of the distribution which weak stationarity requires, rejection of the unit root hypothesis is not sufficient to conclude that the series is stationary.

indexes, the null hypothesis of a single unit root is tested against the alternative of stationarity around a deterministic linear time trend using the following model:

$$\Delta Y_{it} = \alpha + \beta_{il}t + \phi_1 Y_{i,t-1} + \sum_{j=1}^{3} \phi_j \Delta Y_{i,t-j} + \mu_{it} \tag{3}$$

where $\Delta Y_{i,t-j} = Y_{i,t-j} - Y_{i,t-j-1}$, and p is the number of lagged values of first differences (ΔY_{it}). Three lags are included to account for serial correlation in the error terms μ_{it}. If p is constrained to zero, this is simply the Dickey–Fuller test. On the other hand, if lagged values of ΔY_{it} are included ($p \geq 1$) to ensure that the residual μ_{it} is white noise, then this is the augmented Dickey–Fuller test of a single unit root. A time trend is included also in the test because Evans and Savin (1984) showed that pseudo t-statistics are a function of the unknown intercept α. Including a time trend to detrend each series even when the trend coefficient $\beta_{il} = 0$, makes the distribution of the autoregressive parameter estimate ϕ_1 independent of α and thus is probably a prudent decision in performing unit root tests. The empirical percentiles of the test statistics for testing H_o: $\phi_1 = 0$ are tabulated in Fuller (1976, table 8.5.2, p. 373), and improved upon by Mackinnon (1991),[12] to yield more precise critical values for any finite sample size. Using the response surface regression coefficients reported in Mackinnon, the critical value for a 10% (5%) level test of H_o: $\phi_1 = 0$ with a sample size T of 1075 daily price observations are computed as -3.130 (-3.416), with rejection region given by small values (in absolute terms) of the test statistic.

The augmented Dickey–Fuller (ADF) test uses three lags of the first differences (ΔY_{it}). If the null hypothesis of a single unit root is not rejected, a precautionary test of two unit roots is conducted by regressing the second differences in each series against the lagged first difference. An augmented test of two unit roots is similarly undertaken by including up to three lagged values of the second differences. If the null hypothesis of two unit roots is rejected, then the level of the series is declared $I(1)$, or that first differences are $I(0)$. However, as Dickey and Pantula (1987) argued, applying the Dickey–Fuller (DF) test and/or the augmented DF test (ADF) to test for a single unit root makes the asymptotic distributions of the pseudo t statistics sensitive to the number of unit roots present in time series data. In other words, the null hypothesis of a single unit root will be rejected too often. Therefore, a sequential procedure suggested by Dickey and Pantula is used as a complementary test. Essentially, the sequential test proceeds by testing for three, then two, then one unit root in the stock-index and stock-index futures price series.

The results of the above three tests for all four indices are presented in Table I. Panel A shows the results from a test for a single-unit root in the level-specification of each price series, using both the Dickey Fuller (DF) and its Augmented form (ADF) as recommended by Engle and Granger (1987). All t-statistics are negative and well below (in absolute value) their 5% (and even the 10%) critical values so that the null hypothesis

[12]The tabulated critical values for a Dickey–Fuller (or augmented Dickey–Fuller) test for a single unit root as given in Fuller (1976), Engle and Granger (1987), and Engle and Yoo (1987) for example, are based on at most 10,000 replications of Monte Carlo simulations and are given for a limited number of finite sample sizes. One the other hand, critical values given in Mackinnon are based on 25,000 replications, using alternative sample sizes. Asymptotic critical values corresponding to any finite sample size can be easily computed using the tabulated coefficients of the estimated response surface regressions at the 1%, 5%, and 10% levels. Thus, for any sample size T, the estimated critical value is given by

$$\beta_\infty + \beta_1 T^{-1} + \beta_2 T^{-2}$$

where β_∞ denotes estimated asymptotic critical values, and β_1 and β_2 are coefficients on T^{-1}, and T^{-2} in the response surface regression.

Table I

TESTS OF UNIT ROOTS IN STOCK INDEX AND STOCK INDEX FUTURES PRICES

	Pseudo t-statistics							
	Spot Prices				Futures Prices			
Index	**DF**	**DW**	**ADF**	**DW**	**DF**	**DW**	**ADF**	**DW**
	Panel A: Single unit root tests Dickey–Fuller Specification: Spot $\Delta S_t = a + bS_{t-1} + u_t$ Futures: $\Delta F_t = a' + b'F_{t-1} + u'_t$							
S&P 500 Index	–2.63	2.20	–2.87	2.05	–1.53	2.08	–1.65	2.04
FT–SE 100 Index	–1.40	2.15	–2.09	2.06	–1.32	2.23	–2.86	2.02
	Panel B: Two unit root tests Dickey–Fuller Specification: Spot $\Delta^2 S_t = a + b\Delta S_{t-1} + e_t$ Futures: $\Delta^2 F_t = a' + b'F_{t-1} + e'_t$							
S&P 500 Index	–18.93[b]	2.03	–10.75[b]	1.98	–19.76[b]	2.04	–6.95[b]	2.01
FT–SE 100 Index	–18.52[b]	2.09	–10.25[b]	2.00	–15.63[b]	2.19	–6.57[b]	2.01
	Panel C: Dickey–Pantula (DP) sequential tests DP1: $\Delta^3 S_t = c_0 + c_1\Delta^2 S_{t-1} + e_{1t}$ DP2: $\Delta^3 S_t = c'_0 + c'_1\Delta^2 S_{t-1} + c'_2\Delta S_{t-1} + e_{2t}$ DP3: $\Delta^3 S_t = c''_0 + c''_1\Delta^2 S_{t-1} + c''_2\Delta S_{t-1} + c''_3 S_{t-1} + e_{3t}$							

(continued)

Table I (continued)

	Pseudo t-statistics					
	Spot Prices			Futures Prices		
Index	DP1	DP2	DP3	DP1	DP2	DP3
	c1	c2	c3	c1	c2	c3
S&P 500 Index	−23.89[b]	−11.57[b]	−1.14	−21.68[b]	−10.35[b]	−1.57
FT–SE 100 Index	−19.27[b]	−10.67[b]	−2.67	−21.89[b]	−8.36[b]	−2.07

[a]Significance at the 5% level.

[b]Significance at the 1% level.

Notes: Δ_i is the ith difference of spot (S) and Futures (F) prices.

DF: Dickey–Fuller test.

ADF: Augmented Dickey–Fuller test. ADF tests are conducted using up to three lags of the dependent (LHS) variable. In every instance, a maximum of three lagged values of the dependent variable was sufficient to render residuals white noise.

DW: Durbin–Watson statistic.

DP1–DP3: Alternative specifications of the Dickey–Pantula tests. They are illustrated for only the spot indexes to conserve space. Specifications for futures prices are obtained by simply replacing every (S) with (F).

The sample period is 1988:1–1992:5 (1075 observations for each equation).

Critical values for DF and ADF tests are obtained following the procedure in Mackinnon (1991).

of a single unit root is not rejected for any of the four indexes. Evidence from tests of two unit roots using the second-differenced series as the dependent variable is presented in Panel B for all four indexes. Using the DF and ADF tests, the null hypothesis of two unit roots is rejected at 1% level (critical value is −3.9716) for all indexes, as evidenced by the large (in absolute value) t-statistics, thereby providing strong evidence of no unit root in the first differences of each variable. Last but not least, all results from the Dickey–Pantula sequential testing scheme are once more suggestive of only one unit root being present in the level of each series. In sum, only first differencing is needed to render each series stationary. This fulfills the necessary (but not sufficient) condition for a pair of series to be cointegrated—they are integrated of the same order.[13] Therefore, the series are transformed to achieve stationarity by taking the first differences of the natural logarithm of each price. Letting $F(t,T)$ be the price at t of the index futures contract maturing at T, and $S(t)$ be the spot price of the underlying stock index at t, the transformed variables are given by

$$s_t = \ln S(t) - \ln S(t-1)$$

$$f_t = \ln F(t,T) - \ln F(t-1,T)$$

where s_t and f_t represent, respectively, the instantaneous rate of price appreciation in the stock index and the instantaneous relative price change of the futures contract.[14] Panel A in Table II presents the empirical autocorrelation estimates for the transformed spot and futures prices for lags 1 through 12, for both the S&P 500 and the FT–SE 100 indexes. The evidence indicates that the lag one autocorrelation coefficients are reliably negative, and in excess of three standard errors away from zero, for each of the four transformed series, with higher order serial correlations near zero. Unpurged, noise in the transformed price series will tend to reduce the degree of association between the log-transformed spot and futures price-change series. Therefore, the transformed price series is modeled using a moving average process of the first order [i.e., MA(1)]. The innovations are then used as instruments for the cash and futures returns series.[15] Thus,

$$s_t = m + e_t - \theta_1 e_{t-1}$$

and

$$f_t = m' + e'_t - \theta'_1 e_{t-1}$$

where m and m' are the means of the transformed series. Such a process has an autocorrelation function of [see for example, Box and Jenkins (1970)],

$$\rho_k = \begin{cases} \dfrac{-\theta_1}{1+\theta_1^2}, & k = 1 \\ 0, & k > 1 \end{cases}$$

Indeed, a moving average process of the first order is exactly the behavior noted for each of the four transformed series. The value of $\hat{\theta}_1$ for the S&P 500 spot index is 0.459; for the S&P 500 futures index is 0.865; for the FT–SE 100 spot index is 0.855, and the $\hat{\theta}_1$ value for the FT–SE 100 futures is 0.931; all of which satisfy the usual invertibility condition. The larger coefficients of the FT–SE 100 spot and futures series relative to the

[13]If the series are integrated of different orders (e.g., one series is $I(1)$ and the other is $I(2)$]; the two series cannot be cointegrated.

[14]Such stationarity-inducing transformations should not affect the results since, first-differencing and logarithmic transformations are causality-preserving.

[15]This procedure follows Stoll and Whaley (1990) who estimated ARIMA filters to remove noise induced by infrequent trading and bid/ask spreads.

Table II
SAMPLE AUTOCORRELATION COEFFICIENTS OF TRANSFORMED SPOT AND FUTURES PRICES AND THEIR INNOVATIONS GENERATED BY AN MA(1) PROCESS

	S&P 500		FT-SE 100	
Lag	**Spot**	**Futures**	**Spot**	**Futures**
	Panel A: Transformed price series			
1	−0.379[a]	−0.498[a]	−0.495[a]	−0.499[a]
2	0.003	0.008	0.001	0.001
3	0.003	−0.005	−0.002	−0.002
4	−0.009	−0.003	0.004	0.004
5	0.005	0.010	−0.007	−0.008
6	−0.004	−0.010	0.007	0.007
7	0.002	−0.005	−0.003	−0.003
8	−0.005	−0.007	0.006	0.007
9	0.003	−0.011	−0.004	−0.004
10	0.005	0.013	−0.002	0.001
11	−0.004	−0.013	−0.008	−0.004
12	0.000	−0.004	0.000	0.000
Ljung–Box:	273.12[a]	275.54[a]	245.65[a]	300.31[a]
	(0.000)	(0.000)	(0.000)	(0.000)
	Panel B: Price change innovations			
1	0.005	0.008	−0.007	−0.005
2	0.004	0.009	−0.004	0.005
3	−0.002	−0.005	−0.003	−0.003
4	−0.010	−0.004	0.005	0.003
5	−0.006	−0.006	0.000	−0.001
6	−0.008	−0.002	0.012	−0.006
7	−0.002	−0.016	0.010	−0.005
8	−0.011	−0.010	0.014	0.001
9	0.000	0.004	0.005	0.004
10	0.015	0.002	0.004	0.010
11	0.011	0.024	0.003	0.007
12	0.014	0.015	−0.001	0.003
Ljung–Box:	0.99	1.88	0.59	0.32
	(1.000)	(0.999)	(1.000)	(1.000)
Number of observations: 1074				

[a]Denotes significance at the 1% level.

Notes: Ljung Box statistic testing the null hypothesis that all serial correlations taken jointly up to the 12th lag are insignificantly different from zero. The test statistic is distributed as a χ^2 variate with degrees of freedom equal to the number of lags.

Asymptotic standard errors for the autocorrelation coefficients taken individually, can be approximated as the square root of the reciprocal of the number of observations (e.g., ±0.0304 for 1074 observations) under the null hypothesis of zero serial correlation at lag k. One observation is lost when taking first differences.

The sample period is 1988:1–1992:5 (1075 observations for each untransformed series).

S&P 500 series may reflect larger proportional bid–ask spreads for the FT indexes. The innovations (residuals) from each of the fitted MA(1) models are then tested for white noise using the Ljung Box goodness-of-fit test. These results are displayed in panel B in Table II and show that after the effects of the MA(1) process are removed, no serial correlation remains at any lag. Accordingly, a MA(1) process adequately describes the observed structure of the transformed daily series. The transformed series are then used to estimate error correction models and deduce the temporal relationship between stock index and stock index futures prices.

Results From Tests for Cointegration

As mentioned earlier, although each series is found to be nonstationary, there may exist a linear combination of the two series that is stationary. If so, the series are then said to be cointegrated. Because the power of each of the seven tests for cointegration is different, all seven tests are employed. These results are reported in Table III.

The results of the cointegration tests for the S&P 500 spot and futures indexes are presented in panel A. The null of non-cointegration is always reliably rejected at the 1% level using all seven test statistics for the entire period. Sub-period results convey predominantly the same picture, that S&P 500 pair of spot and futures indices are cointegrated. The cointegrating regression parameters for the full period and two sub-periods are also presented. The cointegrating parameters seem to be near unity. Unfortunately, the usual t-test that the slope is 1 is not applicable because of the nonstationarity detected in the levels of each series, which biases the standard errors of the estimated parameters. Similar conclusions can be made regarding cointegration test results involving the pair of spot and futures prices in the U.K., as the evidence presented in panel B is also consistent with the two markets being cointegrated. Accepting that each pair of spot and futures prices forms a cointegrated system, the error correction equation for each index is estimated. The optimal specification of each equation entails a decision about the appropriate number of lags of ΔX_t and ΔY_t. Two modeling approaches are used to arrive at the final specification of the error correction models. Both result in essentially the same models in terms of the root mean squared error (RMSE), the lag-length specification [which is determined using Akaike's Information Criterion (AIC)], and the sign and size of the estimated coefficients. In the first procedure, an Iterative Seemingly Unrelated Regression (ISUR) system of error correction equations is estimated for each pair of spot and futures indexes, with the lag distribution truncated to a maximum of six lags, followed by a series of over- and under-fitting F-tests of each specification, and sequentially eliminating insignificant coefficients as long as this reduces the RMSE and AIC.[16] In the second procedure, each error correction equation is estimated via

[16]Even though a choice of six lags may sound unduly restrictive, it is not unreasonable (and in fact, may be a generous specification) because; (i) informational lags beyond, at most, one day would not be expected to persist if price responses across spot and futures markets conform in a reasonable manner to the assumptions of the efficient market model, and (ii) it is believed that through index arbitrage activities, an effective information link is maintained between spot and futures markets. For the bivariate spot–futures series, Akaike's information criterion is given by:

$$\mathrm{AIC}(p) = \ln \det \sum p + 2d^2p/T \qquad p = 1 \ldots m$$

where d is the dimension of the considered time series, m is the maximum lag order considered, $\sum p$ is the diagonal matrix of sums of squares of residuals generated by ISUR of an AR(p), T is the number of observations used for estimation and det is determinant of a matrix. Since innovations of the transformed variable are used in place of the transformed variables themselves, own lags would not enter significantly into either model—only lags of the other variable in each spot-futures pair is considered in each model. If estimation is done by OLS on equation-by-equation basis, AIC is given by:

$$\mathrm{AIC}(p) = (T - p) \ln \sigma^2 + 2p$$

where all variables are as defined earlier, and σ^2 is the residual variance estimate.

Ordinary Least Squares (OLS), adding up to six lags of the transformed variables, each lag considered one at a time. The lag that further reduces AIC and RMSE below the previous one is added to the model.[17] Both procedures leads to virtually identical error correction model specifications and only one lag is warranted regardless of whether estimation is undertaken on equation-by-equation basis using OLS, or by using a simultaneous estimation technique. Accordingly, the final error correction model for each spot and futures series is given by:

$$s'_t = \alpha_0 - a_1 z_{t-1} + d_1 f'_{t-1} + \mu_{s,t} \tag{4a}$$

$$f'_t = \alpha'_0 - a_2 z_{t-1} + c_1 s'_{t-1} + v_{f,t} \tag{4b}$$

where $z_t = S_t - [b + aF_t]$, and s'_t, and f'_t represent respectively, innovations of the transformed spot and futures price series. Similarly,

$$f'_t = \alpha_1 - a_3 z'_{t-1} + \delta_1 s'_{t-1} + v'_{f',t} \tag{4c}$$

$$s'_t = \alpha'_1 - a_4 z'_{t-1} + \delta'_1 f'_{t-1} + \mu'_{s,t} \tag{4d}$$

where $z'_t = F_t - [b' + a'S_t]$ and f'_t and s'_t are once more the prefiltered innovation series with their X_t and Y_t designations reversed.

Error Correction Model Estimates

OLS parameter estimates of the error correction equations for both the S&P 500 spot and futures innovations are presented in Table IV for the entire period as well as for two sub-periods, while the results for the FT-SE 100 spot and futures innovations are shown in Table V. Four error correction regressions are performed for each pair of spot and futures indexes: (a) spot index innovations are regressed on last period's spot index equilibrium error and last period's futures index innovation; (b) futures index innovations are regressed on last period's spot index equilibrium error and last period's spot index innovation; (c) futures innovations are regressed on last period's futures equilibrium error and spot index innovations; and (d) spot index innovations are regressed on last period's futures index equilibrium error and futures index innovation. These alternative specifications are based on the forward and reverse representations of Engle and Granger's cointegrating regression.

The S&P 500 results presented in panel A of Table IV suggest that all error correction coefficients are reliably different from both zero and unity, at least at the 5% level. This holds true regardless of whether the entire sample or sub-samples are used in the estimation. The full period results (specifications 1 and 2 in panel A) show that subsequent adjustments in spot index values to eliminate equilibrium error in the futures index are rather weak and that spot index adjustments to its own last period's disequilibrium are even weaker. The error correction coefficients on the futures error term is about twice its magnitude on the stock index term (0.188 vs. 0.080) for the entire sample. Thus, about one fifth of last period's futures equilibrium error is eliminated within one day if adjustments are undertaken through price changes in the stock index. On the other hand, about one tenth of last-period's spot index disequilibrium is removed within one day by subsequent spot index adjustments. This differential response is consistent with both lower transaction costs in the futures market and that the spot market is more responsive to shocks in the futures market than to shocks in its own. However, due to

[17]Note that the reported R^2 coefficients are, properly, measures of the between-variables effects, and thus are free of contamination by within-variables effects which arise from including own lagged variables in the regression models [see Pierce (1979)].

Table III

COINTEGRATION TEST RESULTS FOR PAIRS OF SPOT AND FUTURES INDICES

CRDW[a]		DF[b]		ADF[c]		RVAR[d]		ARVAR[e]		UVAR[f]		AUVAR[g]	
S	F	S	F	S	F	S	F	S	F	S	F	S	F
Panel A: S&P 500 Spot-futures tests													
1. Full period (1/4/88–5/30/92)													
1.946[h]	1.881[h]	31.77[h]	66.25[h]	15.204[h]	30.11[h]	306.11[h]	385.21[h]	131.63[h]	202.49[h]	34.45[h]	46.05[h]	25.17[h]	28.96[i]
2. First-half (1/4/88–2/28/90)													
1.735[h]	1.931[h]	20.05[h]	43.37[h]	16.16[h]	19.51[h]	289.16[h]	312.11[h]	127.13[h]	189.61[h]	29.13[h]	30.07[h]	24.69[h]	23.07[i]
3. Second-half (3/1/90–5/30/92)													
1.917[h]	2.377[h]	28.82[h]	29.71[h]	19.53[h]	25.01[h]	279.46[h]	314.91[h]	187.17[h]	192.58[h]	30.93[h]	32.96[h]	26.16[h]	22.61[i]

Cointegrating regressions

1. Full-period

a) $S_t = 43.25 + 0.825F_t$

b) $F_t = 7.66 + 0.867S_t$

2. First-half

c) $S_t = 33.24 + 0.919F_t$

d) $F_t = 18.31 + 0.754S_t$

3. Second-half

e) $S_t = 53.47 + 0.966F_t$

f) $F_t = 3.02 + 1.18S_t$

(continued)

Table III (continued)

CRDW[a]		DF[b]		ADF[c]		RVAR[d]		ARVAR[e]		UVAR[f]		AUVAR[g]	
S	F	S	F	S	F	S	F	S	F	S	F	S	F
Panel B: FT–SE 100 Spot-futures tests													
1. Full period (1/4/88–5/30/92)													
2.131[h]	1.855[h]	43.71[h]	49.27[h]	23.34[h]	43.71[h]	325.48[h]	396.28[h]	148.44[h]	165.01[h]	42.67[h]	56.18[h]	25.63[h]	29.91[i]
2. First-half (1/4/88–2/28/90)													
1.819[h]	1.973[h]	22.11[h]	19.38[h]	13.14[h]	16.59[h]	221.07[h]	256.17[h]	121.39[h]	144.04[h]	20.15[h]	13.56[h]	9.06[h]	11.14[i]
3. Second-half (3/1/90–5/30/92)													
1.896[h]	1.919[h]	33.16[h]	53.18[h]	21.19[h]	25.75[h]	174.04[h]	196.84[h]	122.63[h]	135.83[h]	35.83[h]	27.85[h]	18.85[h]	9.03[i]

Cointegrating Regressions

1. Full-period	2. First-half	3. Second-half
a) $S_t = 174.30 + 0.772F_t$	c) $S_t = 147.01 + 0.716F_t$	e) $S_t = 90.39 + 0.803F_t$
b) $F_t = 113.80 + 0.852S_t$	d) $F_t = 137.77 + 0.757S_t$	f) $F_t = 142.54 + 0.947S_t$

[a]CRDW: Cointegrating Regression Durbin–Watson statistic. Critical values are 0.511 (1%), 0.386 (5%), and 0.322 (10%)—see Engle and Granger ([1987], Table (2)).

[b]DF: Dickey–Fuller test. Critical values are obtained following the method in Mackinnon [1991]—Table (1). For the full period, critical values are 3.909 (1%), 3.343 (5%), and 3.0499 (10%). For the first half period, critical values are 3.9201 (1%), 3.3490 (5%), and 3.0539 (10%). For the second period, critical values are 3.9194 (1%), 3.3486 (5%), and 3.0536 (10%). Critical values for the DF tests are based on 1075 observations (full period), 528 observations (first half), and 547 observations (second half). Minus signs are omitted for simplicity.

[c]ADF: Augmented Dickey–Fuller test. Critical values are same as those used in DF test.

[d]RVAR: Restricted VAR model. Critical values are 18.30 (1%), 13.6 (5%), and 11.0 (10%).

[e]ARVAR: Augmented Restricted VAR model. Critical values are 15.8 (1%), 11.8 (5%), and 9.7 (10%). Test statistic for footnote [d] and [e] is $\tau^2_{\beta_1} + \tau^2_{\beta_2}$.

[f]UVAR: Unrestricted VAR model. Critical values are 23.4 (1%), 18.6 (5%), 16.0 (10%).

[g]AUVAR: Augmented Unrestricted VAR model. Critical values are 22.6 (1%), 17.9 (5%), and 15.5 (10%). Test statistic for footnote [f] and [g] is $2[F_1 + F_2]$.

S and F denote Spot and Futures Price Series, respectively.

[h]Denotes significance at the 1% level.

[i]Denotes significance at the 5% level.

Table IV

OLS PARAMETER ESTIMATES OF THE ERROR CORRECTION MODELS FOR THE S&P 500 SPOT AND FUTURES INNOVATION SERIES

Panel A: LHS variable is S&P 500 Spot Innovations					
Specification 1			**Specification 2**		
$s_t^I = \alpha_0 - a_1 z_{t-1} + d_1 f_{t-1}^I + \mu_{s,t}$			$s_t^I = \alpha_1^I - a_4 z_{t-1}^I + d_1^I f_{t-1}^I + \mu_{s,t}^I$		
$z_t = s_t - (b + aF_t)$			$z_t = F_t - (b' + a'S_t)$		
1. Full-period (1/4/88–5/30/92)			1. Full-period (1/4/88–5/30/92)		
0.180	−0.080[a]	0.111[a]	0.026	−0.159[b]	0.073[b]
(0.409)	(−2.879)	(2.651)	(0.130)	(−2.091)	(2.103)
$R^2 = 0.114$	RMSE = 0.035	$Q(6) =$ 4.89 (0.558)	$R^2 = 0.132$	RMSE = 0.035	$Q(6) =$ 3.55 (0.738)
2. First-half (1/4/88–2/28/90)			2. First-half (1/4/88–2/28/90)		
0.134	−0.121[a]	0.068[b]	0.086	−0.168[a]	0.101[a]
(0.339)	(−3.110)	(2.319)	(0.241)	(−2.717)	(2.648)
$R^2 = 0.081$	RMSE = 0.0651	$Q(6) =$ 3.99 (0.650)	$R^2 = 0.104$	RMSE = 0.051	$Q(6) =$ 2.99 (0.813)
3. Second-half (3/1/90–5/30/92)			3. Second-half (3/1/90–5/30/92)		
0.091	−0.181[a]	0.159[a]	0.034	−0.222[b]	0.156[a]
(0.665)	(−4.853)	(3.436)	(0.099)	(−2.445)	(2.499)
$R^2 = 0.149$	RMSE = 0.049	$Q(6) =$ 1.76 (0.880)	$R^2 = 0.150$	RMSE = 0.040	$Q(6) =$ 1.83 (0.778)

(continued)

Table IV (continued)

Panel B: LHS variable is S&P 500 futures innovations

Specification 1			Specification 2		
$f_t' = \alpha_0' - a_2 z_{t-1} + c_1 s_{t-1}' + v_{f,t}$ $z_t = s_t - (b + aF_t)$			$f_t' = \alpha_1 - a_3 z_{t-1}' + d_1 s_{t-1}' + v_{f,t}'$ $z_t' = F_t - (b' + a'S_t)$		
1. Full-period (1/4/88–5/30/92)			1. Full-period (1/4/88–5/30/92)		
0.092	−0.653[a]	0.236[a]	0.168	−0.049[b]	0.147[b]
(0.045)	(−4.915)	(3.447)	(0.091)	(−2.061)	(2.212)
$R^2 = 0.484$	RMSE = 0.011	$Q(6) =$ 1.39 (0.781)	$R^2 = 0.103$	RMSE = 0.039	$Q(6) =$ 0.99 (0.931)
2. First-half (1/4/88–2/28/90)			2. First-half (1/4/88–2/28/90)		
0.108	−0.844[a]	0.281[a]	0.044	−0.051[b]	0.257[a]
(0.536)	(−5.410)	(3.933)	(0.010)	(−2.348)	(3.386)
$R^2 = 0.318$	RMSE = 0.018	$Q(6) =$ 2.16 (0.880)	$R^2 = 0.099$	RMSE = 0.045	$Q(6) =$ 2.77 (0.761)
3. Second-half (3/1/90–5/30/92)			3. Second-half (3/1/90–5/30/92)		
0.215	−0.898[a]	0.307[a]	0.016	−0.134[b]	0.312[a]
(0.767)	(−5.341)	(4.917)	(0.011)	(−2.565)	(3.431)
$R^2 = 0.456$	RMSE = 0.013	$Q(6) =$ 2.13 (0.871)	$R^2 = 0.179$	RMSE = 0.036	$Q(6) =$ 1.88 (0.790)

[a]Denotes significance at the 1% level.

[b]Denotes significance at the 5% level.

Notes: z_{t-1} and z_{t-1}' are the one-period lagged values of the error correction terms, representing the unexpected components of the spot and futures level series, respectively.

S_t and F_t are spot and futures prices in levels, respectively.

s_t and f_t' are spot and futures innovations, respectively.

RMSE is Root Mean Square Error.

Q(6) is the Ljung–Box goodness-of-fit statistic; testing the null hypothesis that all serial correlations up to the sixth lag are jointly zero. Numbers in parentheses are marginal significance levels of the Q-statistics.

Other numbers in parentheses are t-statistics.

the higher cost-of-adjustment in the spot market, only a minor portion of the equilibrium error is eliminated within a day. In any event, full adjustment to long-run equilibrium is never completed within one day. Another phenomenon of interest is the spot market's response to last period's futures innovation. Once adjustments to the equilibrium error are controlled for, although weak, a statistically significant stock index adjustment is still detected, regardless of the cointegrating regression specification used. This may be indicative of weak causality running from futures-to-spot. Furthermore, causality from futures-to-spot appears to have strengthened over time. The causality coefficients of the second half are more than double those of the first half. When coupled with a trend towards a greater speed of adjustment over time, the evidence indicates closer integration of the spot and futures markets. Perhaps this is a result of more active index arbitrage and more efficient program trading. In all sample periods, the Ljung–Box χ^2 statistics attest to the goodness-of-fit of the error correction models.

Panel B in Table IV presents the results from estimating an error correction model with futures innovations as the LHS variable. Once more, two alternative cointegrating regression specifications are used. The first uses the spot index equilibrium error as the error correction term (specification 1 in panel B), while the second uses futures index disequilibrium as the error correction term (specification 2 in panel B). More interesting phenomena are shown in specification (1) of panel B, the most important of which is the pronounced leading effect of stock market. Specifically, the magnitudes of the error correction coefficients are now stronger and closer to unity, compared to their magnitudes reported in panel A. However, they are still reliably different from unity. This is true for the entire sample and, also, over the two sub-samples. It appears that spot index mispricing triggers subsequent price reaction in the futures market, more so, than mispricing in the latter motivating price changes in the former. The major component of adjustment towards equilibrium spot-futures pricing, however, appears to be accomplished primarily through changes in futures, rather than spot prices. This is hardly surprising considering the well established transaction cost-advantage of futures markets. Second, the coefficients on the one-day lag spot innovations are more than twice their magnitude on the one-day lag futures innovations [specification (2) of panel A]. Together, these results indicate stronger causal forces flowing from spot-to-futures, more so than the reverse.

The results of specification (2) in panel B are based on futures equilibrium errors as the error correction term, and serve to confirm the stronger lead from spot-to-futures. A cursory comparison of the R^2 coefficients in Table IV suggest that specifications (2) and (1) of panels A and B, respectively, are better representations of the temporal relation between spot and futures markets. In summary, the results in Table IV suggest that feedback exists between the two markets which may be attributed to both short-run dynamics and adjustments to long-run equilibrium pricing. The relative magnitudes of the error correction coefficients shown in panels A and B of Table IV indicate that (i) adjustments to equilibrium spot-futures pricing are accomplished primarily through subsequent futures as opposed to spot index price changes, and (ii) futures prices are more responsive to lagged spot index equilibrium pricing errors and spot index innovations than are spot prices to lagged futures pricing errors and futures innovations. The lead from spot-to-futures is probably stronger when viewed relative to the lead from futures-to-spot, on a daily basis.

This evidence is in contrast with Ng's (1987) results, who also used daily data, and reported evidence which supports a tendency of the futures market to lead the spot market. This may perhaps, be attributed to the dynamics of the equilibrium error, which are explicitly captured by the error correction model. On the other hand, this evidence is in line with evidence reported in Chan, Chan, and Karolyi (1991), who found stronger spot-to-futures effects than was found in previous studies, when the lead–lag model is specified to control for relevant omitted factors.

For the sake of completeness, the strength of the lead–lag relation when including contemporaneous values of the endogenous variables is investigated also. Specifically, the one-day lag coefficient in each error correction equation is constrained to zero, allowing only the contemporaneous, and the error correction coefficients to remain unconstrained in eqs. (4a)–(4d). Therefore, eqs. (4a) and (4d) are respecified to include contemporaneous futures innovations, while eqs. (4b) and (4c) are respecified to include contemporaneous spot innovations.[18] This introduces simultaneity in the estimation process which makes parameter estimates biased and inconsistent. To overcome this problem, the extended equations are estimated using Three Stage Least Squares which potentially produces unbiased, consistent, and efficient parameter estimates. All contemporaneous coefficients are close to (but statistically different from) unity. Furthermore, linear restrictions (using an F-test) on the one-day lag coefficients of the endogenous variables are always rejected with almost certainty (not reported). This serves to confirm that both spot and futures indexes are related on a predominantly instantaneous basis but that reliably, some lagged effects also prevail.

Because futures price variability may be higher than the rest of the sample as the contract nears maturity,[19] [see for example, Milonas (1986)] it is of interest to examine whether the observed lead–lag patterns are affected by some abnormal price behavior close to expiration days. Therefore, the last 2 weeks of each quarterly contract period are excluded, and the error correction models are reestimated using the spliced series.[20] Although slightly different parameter estimates are obtained, the results and conclusions are virtually unchanged (not reported) for both the S&P 500 and FT–SE 100 pairs of indexes. It may be that futures volatility in the last 2 weeks is not any different from its level in other days prior to expiration; or that, if it is different, it does not appear to alter the temporal causal patterns in any statistically meaningful manner when using daily data.

Table V presents the parameter estimates of the error correction models for the FT–SE 100 spot and futures prices using the two alternative specifications of the cointegrating regressions. These results are very similar to the S&P 500 spot-futures relationship. First, the relative magnitudes of the error correction coefficients across panels A and B suggest that feedback is present but that the futures response coefficients [specification (1) in panel B] are more than twice as strong as the spot market's response coefficients [specification (2) in panel A]. Not only are such coefficients significantly different from zero at the 1% level, they are also significantly different from unity. However, they possess their expected sign. Second, further evidence supporting feedback is apparent from examining the statistical significance of the lagged coefficients on the spot and futures innovations. Both a lead from spot-to-futures [specification (1) in panel B] and a lead from futures-to-spot [specification (2) in panel A] are detected. These coefficients are reliably non-zero at the 1% level for almost all of the sample periods

[18] Respecifying error correction models to include contemporaneous endogenous variables is not an uncommon practice [see, for e.g., Hall (1986)].

[19] See, for example, Milonas (1986) for some empirical results on the relation between price variablility and contract maturity.

[20] A switch over to the next maturing futures contract is essentially assumed to take place 2 weeks before maturity of each quarterly contract so that arbitrageurs reverse their positions around that day. Results after splicing are obtainable from the authors upon request.

Table V

OLS PARAMETER ESTIMATES OF THE ERROR CORRECTION MODELS FOR THE FT–SE 500 SPOT AND FUTURES INNOVATION SERIES

Panel A: LHS variable is FT–SE 500 Spot Innovations

Specification 1			**Specification 2**		
$s'_t = \alpha_0 - a_1 z_{t-1} + d_1 f'_{t-1} + \mu_{s,t}$			$s'_t = \alpha'_1 - a_4 z'_{t-1} + d'_1 f'_{t-1} + \mu'_{s,t}$		
$z_t = S_t - (b + aF_t)$			$z'_t = F_t - (b' + a'S_t)$		
1. Full-period (1/4/88–5/30/92)			1. Full-period (1/4/88–5/30/92)		
0.052	−0.072[a]	0.067[b]	0.064	−0.266[a]	0.187[a]
(0.025)	(−2.761)	(2.122)	(0.030)	(−2.511)	(2.109)
$R^2 = 0.091$	RMSE = 0.066	$Q(6) = 3.19$	$R^2 = 0.198$	RMSE = 0.036	$Q(6) = 1.79$
		(0.785)			(0.938)
2. First-half (1/4/88–2/28/90)			2. First-half (1/4/88–2/28/90)		
0.172	−0.117[a]	0.072[b]	0.031	−0.218[b]	0.079[b]
(0.135)	(−2.837)	(2.708)	(0.009)	(−2.122)	(2.213)
$R^2 = 0.158$	RMSE = 0.041	$Q(6) = 1.78$	$R^2 = 0.117$	RMSE = 0.045	$Q(6) = 1.33$
		(0.833)			(0.767)
3. Second-half (3/1/90–5/30/92)			3. Second-half (3/1/90–5/30/92)		
0.088	−0.099[a]	0.114[b]	0.116	−0.247[a]	0.150[a]
(0.113)	(−2.767)	(2.321)	(0.413)	(−3.192)	(2.997)
$R^2 = 0.177$	RMSE = 0.038	$Q(6) = 2.12$	$R^2 = 0.212$	RMSE = 0.031	$Q(6) = 0.81$
		(0.808)			(0.936)

(continued)

Table V (continued)

Panel B: LHS variable is FT–SE 500 futures innovations

Specification 1			**Specification 2**		
$f_t' = \alpha_0' - a_2 z_{t-1} + c_1 s_{t-1}' + v_{f,t}$			$f_t' = \alpha_1 - a_3 z_{t-1}' + d_1 s_{t-1}' + v_{f,t}'$		
$z_t = s_t - (b + aF_t)$			$z_t' = F_t - (b' + a'S_t)$		
1. Full-period (1/4/88–5/30/92)			1. Full-period (1/4/88–5/30/92)		
0.068	−0.693[a]	0.284[a]	0.095	−0.170[a]	0.261[a]
(0.090)	(−3.628)	(3.538)	(0.450)	(−2.836)	(3.540)
$R^2 = 0.222$	RMSE = 0.017	$Q(6) = 2.69$ (0.846)	$R^2 = 0.133$	RMSE = 0.047	$Q(6) = 1.78$ (0.912)
2. First-half (1/4/88–2/28/90)			2. First-half (1/4/88–2/28/90)		
0.109	−0.505[a]	0.144[a]	0.218	−0.099[a]	0.191[a]
(0.212)	(−3.113)	(2.656)	(0.916)	(−2.731)	(2.996)
$R^2 = 0.191$	RMSE = 0.023	$Q(6) = 1.77$ (0.662)	$R^2 = 0.149$	RMSE = 0.033	$Q(6) = 1.13$ (0.889)
3. Second-half (3/1/90–5/30/92)			3. Second-half (3/1/90–5/30/92)		
0.079	−0.718[a]	0.269[a]	0.166	−0.213[a]	0.241[a]
(0.118)	(−4.298)	(3.113)	(0.576)	(−2.891)	(2.771)
$R^2 = 0.241$	RMSE = 0.019	$Q(6) = 0.89$ (0.919)	$R^2 = 0.186$	RMSE = 0.026	$Q(6) = 1.34$ (0.767)

[a]Denotes significance at the 1% level.
[b]Denotes significance at the 5% level.

Notes: z_{t-1} and z_{t-1}' are the one-period lagged values of the error correction terms, representing the unexpected components of the spot and futures level series, respectively.

S_t and F_t are spot and futures prices in levels, respectively.

s_t and f_t' are spot and futures innovations, respectively.

RMSE is Root Mean Square Error.

Q(6) is the Ljung–Box goodness-of-fit statistic; testing the null hypothesis that all serial correlations up to the sixth lag are jointly zero. Numbers in parentheses are marginal significance levels of the Q-statistics.

Other numbers in parentheses are t-statistics.

investigated. As in the case of the S&P 500, the causal coefficients in specification (1) in panel B are about twice their magnitude reported in specification (2) in panel A, which again suggests that the FT–SE 100 spot market lead is probably stronger relative to the FT–SE 100 futures lead. The Ljung–Box statistics are also suggestive of specification-adequacy. Tests for insignificant lagged effects, once contemporaneous effects are controlled for, reject the null of no lagged responses at the very stringent 0.01% level (not reported). Additionally, using a similar procedure, expiration-day effects if any, do not affect the basic lead–lag results and conclusions.

To summarize, the conclusions from Tables IV and V are remarkably similar. Feedback is detected between spot and futures markets so that no single market can be considered weakly exogenous. Equilibrium adjustments are not fully completed within one day and contemporaneous linkages are, obviously, not the sole medium of interaction between stock index and stock index futures markets. However, it can be argued also that presence of causal linkages (and absence of completely contemporaneous linkages) is consistent with a variety of market imperfections and realities that prohibit market participants from responding to every single deviation between actual and theoretically correct futures value estimates. In essence, what matters is whether such delayed responses are large enough to generate economic profits, transaction, and other costs considered.

Tests of Parameter Stability and Quarterly Lead–Lag Results

A test of parameter stability is warranted in light of the evidence presented in Tables IV and V where parameter estimates are suspected to be temporally unstable and are dependent on the specification used. Therefore, the Farely–Hinich test [Farely, Hinich, and McGuire (1975)] for parameter stability is undertaken. In the Farley–Hinich test, each of the coefficients which is suspected of instability is modeled as a linear continuous function of time, and the significance of the newly created coefficients is checked.[21] Since specifications (2) (of panel A) and (1) (of panel B) are uniformly the best representations (with the highest R^2) for the S&P 500 and the FT–SE 100 indexes, the tests are restricted to only those two cases. The significance of the a_4', δ_1'', a_2', and c_1' in the following extended versions of eqs. (4d) and (4b), respectively; as given in eqs. (5a) and (5b), is tested

$$s_t' = \alpha_0 - a_4 z_{t-1}' + \delta_1' f_{t-1}' + a_4' T z_{t-1}' + \delta_1'' T f_{t-1}' + \mu_{s,t}'' \tag{5a}$$

$$f_t' = \alpha_0' - a_2 z_{t-1} + c_1 s_{t-1}' + a_2' T z_{t-1} + c_1' T s_{t-1}' + v_{f,t}'' \tag{5b}$$

where T is a linear time trend. If a_4' and a_2' are significantly different from zero, the error correction coefficients in the spot and futures innovation equations are declared time varying. Likewise, the hypotheses that the causal coefficients are temporally stable will be rejected if the δ_1'' and c_1' coefficients are statistically significant at conventional levels. The Farley–Hinich test results for both S&P 500 and FT–SE 100 spot-futures pairs are reported in Table VI.

Panels A and B present the results for the S&P 500 pair while panels C and D present these results for the FT–SE 100 pair. As is apparent, structural changes in the causal coefficients δ_1'' and c_1', but not in the error correction coefficients a_4' and a_2', is confirmed for the S&P 500 pair. the FT–SE 100 results indicate temporal changes in both the error

[21]This, by no means, is the only possible behavior of the estimated parameters. Besides using a deterministic linear trend model, one can think of a variety of possible stochastic parameter models.

Table VI
FARLEY–HINICH TEST OF PARAMETER STABILITY CONDUCTED WITH THE ERROR CORRECTION EQS. (5A) AND (5B) OVER THE FULL SAMPLE PERIOD (1/4/88–5/30/92)

Panel A: LHS Variable is S&P 500 Spot Price Change Innovations

Specification:
$s'_t = \alpha_0 - a_4 z'_{t-1} + \delta'_1 f'_{t-1} + a'_4 T z'_{t-1} + \delta'' T f'_{t-1} + \mu''_{s,t}$
$z_t = F_t - (b' + a' S_t)$

Intercept	a_4	δ'_1	a'_4	δ''_1	R^2	DW
0.0115	−0.166[a]	0.092[b]	0.0091	0.0238[b]	0.186	2.035
(0.093)	(−2.517)	(2.219)	(0.753)	(2.420)		

Panel B: LHS variable S&P 500 futures price change innovations

Specification:
$f'_t = \alpha'_0 - a_2 z_{t-1} + c_1 s'_{t-1} + a'_2 T z_{t-1} + c'_1 T s'_{t-1} + v''_{f,t}$
$z_t = S_t - (b + aF_t)$

Intercept	a_2	c_1	a'_2	c'_1	R^2	DW
0.176	−0.822[a]	0.262[a]	0.0019	0.0771[b]	0.501	1.990
(0.312)	(5.313)	(2.996)	(1.101)	(2.096)		

Panel C: LHS variable is FT–SE 100 spot price change innovations

Specification:
$s'_t = \alpha_0 - a_4 z'_{t-1} + \delta'_1 f'_{t-1} + a'_4 T z'_{t-1} + \delta'' T f'_{t-1} + \mu''_{s,t}$
$z_t = F_t - (b' + a' S_t)$

Intercept	a_4	δ'_1	a'_4	δ''_1	R^2	DW
0.0716	−0.192[a]	0.133[a]	0.0153[b]	0.0197[b]	0.266	1.989
(0.113)	(−2.446)	(2.695)	(2.117)	(2.168)		

Panel D: LHS variable is FT–SE 100 futures price change innovations

Specification:
$f'_t = \alpha'_0 - a_2 z_{t-1} + c_1 s'_{t-1} + a'_2 T z_{t-1} + c'_1 T s'_{t-1} + v''_{f,t}$
$z_t = S_t - (b + aF_t)$

Intercept	a_2	c_1	a'_2	c'_1	R^2	DW
0.5856	−0.565[a]	0.263[a]	0.053[b]	0.133[b]	0.298	2.110
(0.357)	(−3.167)	(2.991)	(2.559)	(2.111)		

[a]Indicates statistical significance at the 1% level.
[b]Indicate statistical significance at the 5% level, respectively.
Notes: Numbers reported below in parentheses are t-statistics.
DW: is Durbin–Watson statistic.

correction and causal coefficients for both spot and futures innovations. It is also noted that all of the statistically significant trend coefficients in panels A through D have positive signs, revealing that, in addition to being unstable, the slopes are, on average, trending upwards over time. Having rejected coefficient stability over time, it would be illuminating to reexamine the lead–lag relations over shorter periods. Therefore, only eqs. (4b) and (4d) are reestimated over each quarterly period. These results are reported in Table VII for both the S&P 500 and FT–SE 100 pairs of indexes, and suggest that, indeed, many of the coefficients do not appear to be stable. Statistical reference based on a constant coefficient assumption, using the full sample in estimation, can be misleading.

Out-Of-Sample Fit and Forecast Evaluation

To account for the fact that the parameters are not constant over time and to further check the validity of the error correction specification of the spot-futures lead–lag relation, the ability of these regressions to make out-of-sample multistep-ahead forecasts one quarter at a time is investigated. Thus, in each quarter, error correction model parameters are estimated and are then used to produce forecasts over the following quarter. With 18 quarters in this study period, parameter estimates are sequentially updated on a quarterly basis, before a forecast over the subsequent quarter is generated. These forecasts are then compared with forecasts based on a standard vector autoregression (VAR) model[22] which is frequently used [see for example, Kwaller et al. (1987)] in similar investigations of the lead–lag relation. Accuracy of multistep-ahead forecasts is judged by the mean absolute forecast error (MAE) given by

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\hat{Y}_{T+i} - Y_{T+i}| \tag{6}$$

where $\hat{Y}_{T+i}$ is an i-step-ahead forecast of Y_{T+i}, n is the number of trading days in any quarterly forecast period, and T is the terminal day in the estimation period. To conserve space, out-of-sample forecasting results based on only two of the four error correction specifications, notably eqs. (4b) and (4d) are reported, together with the VAR forecast results. These results are given in Table VIII for both the S&P 500 and FT–SE 100 pairs of indexes. The evidence is broadly consistent with the superiority of forecasts produced by an error correction model in relation to those generated from a standard VAR model. In many cases the MAE statistics are cut in half when using an error correction, as opposed to a VAR model; a result that is to be expected since the error correction coefficients are generally statistically significant in the error correction representations. Ignoring the error correction term may lead to a misspecified lead–lag spot–futures relation. Stated alternatively, including an error correction term should lead to forecasts that are at least as good as (if not better than) VAR-based forecasts.

Error Correction Model Diagnostics

The final error correction eqs. (4b) and (4d) are checked for statistical adequacy. Ordinary Least Squares estimation assumes that residuals are normal, conditionally homoskedastic, and mutually serially uncorrelated at all lags. If errors are autocorrelated or serially heteroskedastic, OLS estimation yields biased estimates of the parameters variance–covariance matrix. If the residuals are not normally distributed, parameter estimates are no longer efficient, and conventional t tests are not necessarily valid in finite samples. Normality is checked using the Shapiro–Wilk statistic [Judge et al., (1985, p. 886)]. Although not reported, the values of skewness and the Shapiro–Wilk statistics are consistent with the normal distribution. Individual and joint tests of first and higher

[22]The vector autoregression model employed uses spot and futures price change innovations, and is specified identically to the error correction models used, except that no error correction term is included in any of the equations. This allows for an examination of the marginal contribution of the lagged error correction term for enhancing prediction accuracy.

order serial correlation in the residuals are conducted using the Ljung–Box Q statistic. These results are reported earlier in Tables IV through V, and are generally consistent with statistical adequacy of the adopted lag specification. Conditional homoskedasticity is tested by regressing the squared residuals from each error correction specification onto a constant and six lagged values. The null hypothesis is that of conditional homoskedasticity which would imply that coefficients on the lagged squared residuals should be jointly insignificantly different from zero, using an F-statistic. The null hypothesis is not rejected at the 10% level. Therefore, the assumptions seem reasonably tenable.

SUMMARY AND CONCLUDING REMARKS

Cointegration analysis is used to examine the temporal causal linkage between stock index and stock index futures prices for both the S&P 500 and FT–SE 100 indexes over the period from 1988 to 1992. The findings are summarized as follows. First, the cash and futures markets are cointegrated so that an error correction representation for each series is appropriate. Second, the results are generally consistent with market efficiency. Spot and futures prices appear to be mostly simultaneously related on a daily basis and that lagged interactions, although statistically significant, are rather weak in a magnitude so that their

Table VII
QUARTER-BY-QUARTER ESTIMATION OF THE ERROR CORRECTION EQS. (4B) AND (4D) FOR BOTH THE S&P 500 AND FT–SE 100 SPOT AND FUTURES PRICE-CHANGE INNOVATIONS

	Specification Eq. (4b)				Specification: Eq. (4d)			
Quarter	Intercept	z_{t-1}	s'_{t-1}	R^2	Intercept	z_{t-1}	f'_{t-1}	R^2
	Panel A: S&P 500 price change innovations							
Q1-1988	0.0184	−0.709[a]	0.284[a]	0.371	0.0027	−0.078[a]	0.111[a]	0.083
Q2-1988	0.0316	−0.688[a]	0.287[a]	0.379	0.0034	−0.035	0.127[b]	0.169
Q3-1988	0.0066	−0.483[b]	0.354[a]	0.473	0.0011	−0.174[a]	0.229[a]	0.440
Q4-1988	0.0118	−0.493[b]	0.272[a]	0.414	0.0006	−0.153[a]	0.371[b]	0.475
Q1-1989	0.0143	−0.633[a]	0.373[a]	0.387	0.0021	−0.054	0.161[b]	0.286
Q2-1989	0.0102	−0.503[a]	0.391[a]	0.474	0.0021	−0.166[b]	0.245[a]	0.252
Q3-1989	0.0225	−0.119[b]	0.389[a]	0.299	0.0037	−0.127[a]	0.412[a]	0.333
Q4-1989	0.0323	−0.353[b]	0.343[a]	0.252	0.0062	−0.227[a]	0.296[b]	0.281
Q1-1990	0.0444	−0.384[b]	0.224[b]	0.411	0.0001	−0.035	0.122[b]	0.049
Q2-1990	0.0099	−0.412[b]	0.404[a]	0.463	0.0027	−0.129[b]	0.118[b]	0.188
Q3-1990	0.0321	−0.162[b]	0.331[a]	0.479	0.0092	−0.058	0.217[a]	0.059
Q4-1990	0.0255	−0.046	0.293[a]	0.273	0.0025	−0.279[a]	0.104[b]	0.244
Q1-1991	0.0078	−0.241[b]	0.248[b]	0.161	0.0011	−0.533[a]	0.131[b]	0.096
Q2-1991	0.0015	−0.748[a]	0.172[b]	0.213	0.0017	−0.213[b]	0.128[b]	0.089
Q3-1991	0.0169	−0.993[a]	0.207[b]	0.250	0.0015	−0.047	0.265[a]	0.101
Q4-1991	0.0354	−0.760[a]	0.156[b]	0.183	0.0071	−0.201[b]	0.193[a]	0.329
Q1-1992	0.0138	−0.801[a]	0.274[b]	0.243	0.0063	−0.201[b]	0.209[b]	0.229
Q2-1992	0.0072	−0.759[a]	0.373[b]	0.265	0.0045	−0.153[b]	0.204[b]	0.285

Table VII (continued)

Quarter	Specification Eq. (4b) Intercept	z_{t-1}	s'_{t-1}	R^2	Specification: Eq. (4d) Intercept	z_{t-1}	f'_{t-1}	R^2
	Panel B: FT–SE 100 price change innovations							
Q1-1988	0.0216	–0.323[b]	0.478[a]	0.262	0.0064	–0.178[b]	0.241[b]	0.202
Q2-1988	0.0069	–0.255[b]	0.235[a]	0.383	0.0044	–0.099	0.047	0.017
Q3-1988	0.0049	–0.532[a]	0.123[a]	0.244	0.0069	–0.053	0.338[a]	0.310
Q4-1988	0.0028	–0.368[b]	0.163[b]	0.441	0.0028	–0.116[b]	0.229[b]	0.200
Q1-1989	0.0020	–0.423[a]	0.243[b]	0.477	0.0070	–0.144[b]	0.096	0.133
Q2-1989	0.0061	–0.126[b]	0.250[a]	0.157	0.0045	–0.105[b]	0.148[b]	0.202
Q3-1989	0.0026	–0.298[a]	0.197[b]	0.226	0.0005	–0.215[b]	0.163[b]	0.254
Q4-1989	0.0079	–0.077	0.017	0.035	0.0002	–0.397[b]	0.204[b]	0.367
Q1-1990	0.0060	–0.109[b]	0.121[b]	0.159	0.0002	–0.307[b]	0.077	0.277
Q2-1990	0.0011	–0.982[a]	0.147[b]	0.481	0.0009	–0.588[b]	0.136[b]	0.249
Q3-1990	0.0077	–1.081[a]	0.232[a]	0.331	0.000	–0.221[b]	0.167[b]	0.246
Q4-1990	0.0018	–0.730[a]	0.142[b]	0.205	0.0010	–0.565[a]	0.143[b]	0.169
Q1-1991	0.0150	–1.068[a]	0.389[a]	0.259	0.0057	–0.296[b]	0.143[b]	0.140
Q2-1991	0.0015	–0.822[a]	0.109[b]	0.403	0.0022	–0.217[b]	0.150[b]	0.357
Q3-1991	0.0083	–0.515[a]	0.132[b]	0.191	0.0029	–0.431[a]	0.179[b]	0.208
Q4-1991	0.0069	–0.634[a]	0.153[b]	0.301	0.0028	–0.331[b]	0.179[b]	0.219
Q1-1992	0.0042	–0.791[a]	0.225[a]	0.374	0.0059	–0.584[a]	0.123[b]	0.266
Q2-1992	0.0045	–0.705[a]	0.105[b]	0.320	0.0003	–0.166[b]	0.241[b]	0.291

[a]Denotes statistical significance at the 1% level.

[b]Denotes statistical significance at the 5% level.

Notes: Eqs. (4b) and (4d) were jointly estimated using Generalized Least Squares. The error correction term in eq. (4b) represents the spot index price level shock, while the error correction term in eq. (4d) represents the futures price level shock.

All intercept coefficients were statistically insignificant at the 20% level.

The spot and futures variables are transformed price-change innovations.

predictive content may not be economically significant. Third, the relative magnitudes of the error correction and lagged coefficients suggest that although feedback exists between the cash and futures markets for both the S&P 500 and FT–SE 100 indexes, the spot-to-futures lead appears to be more pronounced across days relative to the futures-to-spot lead. Fourth, futures prices exhibit stronger subsequent responses to disequilibrium in spot prices that do spot prices to last period's futures equilibrium error. These results are generally robust to alternative error correction specifications, and time periods examined. Fifth, judging from the out-of-sample tests, the proposed error correction specification achieves a significantly lower MAE in the majority of cases when compared with the forecasting performance of a standard vector autoregression model. Indeed, as pointed out by Granger (1986), an error correction model should produce better short-run forecasts and will certainly produce long-run forecasts that hold together in an economically meaningful way. The evidence is, therefore, consistent with the important price discovery role served by both stock index and stock index futures markets, and corroborates the findings of Chan et al. (1991) of a much stronger interdependence between stock index cash and futures markets than found in earlier studies.

Table VIII
PERFORMANCE OF THE ONE-QUARTER-AHEAD FORECASTS BASED ON THE ERROR CORRECTION MODELS (ECM) OF EQS. (4B) AND (4D), AND VECTOR AUTOREGRESSION VAR MODELS. FORECAST EVALUATION IS FOR THE PERIOD: Q2–1988 TO Q2–1992

	S&P 500				FT–SE 100			
	ECM		VAR		ECM		VAR	
	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures
Quarter	MAE	MAE	MAE	MAE	MAE	MAE	MAE	MAE
Q2:88	0.778	0.837	1.656	1.956	0.735	0.688	1.309	0.735
Q3:88	0.590	0.994	0.648	1.001	0.626	0.371	0.732	0.747
Q4:88	0.277	0.507	0.507	0.821	0.149	0.626	0.208	0.702
Q1:89	0.174	0.406	0.261	0.511	0.871	0.170	1.023	0.258
Q2:89	0.236	0.468	0.309	0.657	0.960	0.945	1.030	1.088
Q3:89	0.274	0.704	0.350	0.734	0.327	0.842	0.531	0.969
Q4:89	0.380	0.592	0.472	1.070	0.234	0.351	0.288	0.417
Q1:90	0.122	0.519	0.437	0.871	0.157	0.159	0.266	0.240
Q2:90	0.127	0.259	0.195	0.333	0.194	0.144	0.292	0.272
Q3:90	0.082	0.661	0.187	0.711	0.472	0.215	0.620	0.291
Q4:90	0.739	0.797	0.797	0.858	0.412	0.735	0.502	0.753
Q1:91	0.616	0.265	0.726	0.426	0.884	0.625	0.952	0.710
Q2:91	0.080	0.099	0.111	0.266	0.596	0.706	0.668	0.999
Q3:91	0.096	0.997	0.188	1.066	0.444	0.789	0.495	1.011
Q4:91	0.223	0.182	0.277	0.286	0.289	0.440	0.496	0.584
Q1:92	0.128	0.341	0.190	0.663	0.017	0.018	0.148	0.145
Q2:92	0.118	0.128	0.196	0.151	0.241	0.176	0.295	0.260

Notes: The MAEs were calculated from the log-transformed price change innovations.
MAE denotes the Mean Absolute Error.

Bibliography

Akaike, H. (1969a): "Statistical Predictor Identification" *Annals of the Institute of Statistical Mathematics,* 21:203–217.

Akaike, H. (1969): "Fitting Autoregressions for Prediction," *Annals of the Institute of Statistical Mathematics,* 21:243–247.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics,* 31:307–327.

Box, G. E. P., and Jenkins, G. M. (1976): *Time Series Analysis: Forecasting and Control,* 2nd ed. San Francisco: Holden-Day.

Chan, K., Chan, K. C., and Karolyi, G. A. (1991): "Intraday Volatility in the Stock Index and Stock Index Futures Markets," *Review of Financial Studies,* 4:657–683.

Cornell, B. and French, K. (1983, Spring): "The Pricing of Stock Index Futures," *The Journal of Futures Markets,* 3:1–14.

Cornell, B., and French, K. (1983, June): "Taxes and the Pricing of Stock Index Futures," *Journal of Finance,* 38:675–694.

Dickey, D. A., and Fuller, W. A. (1979, June): "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association*, 74:427–431.

Dickey, D. A., and Fuller, W. A. (1981, July): "Likelihood Ratio Statistics For Autoregressive Time Series With a Unit Root," *Econometrica*, 49:1057–1072.

Dickey, D. A., Fuller, W. A., and Pantula, S. (1987, October): "Determining the Order of Differencing in Autoregressive Processes," *Journal of Business and Economic Statistics*, 4:455–461.

Engle, R. B., and Granger, C. W. (1987): "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55:251–276.

Engle, R. B., and Yoo, B. S. (1987): "Forecasting and Testing in Cointegrated Systems," *Journal of Econometrics*, 35:143–159.

Engle, R. B., and Yoo, B.-S. (1982): "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation," *Econometrica*, 50:987–1008.

Evans, G. and Savin, N. (1984): "Testing for Unit Roots: 2," *Econometrica*, 52:1241–1269.

Farley, J. U., and Hinich, M. J. (1970): "A Test For a Shifting Slope Coefficient in a Linear Model," *Journal of the American Statistical Association*, 65:1320–1329.

Farley, J. U., and T. W. McGuire (1975): "Some Comparisons of Tests for A Shift in the Slopes of a Multivariate Linear Time Series Model," *Journal of Econometrics*, 3:297–318, 379–393.

Figlewski, S. (1984, July): "Hedging Performance and Basis Risk in Stock Index Futures," *Journal of Finance*, 39:657–669.

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

Granger, C. (1969): "Investigating Causal Relations by Econometric Models and Cross Spectral Methods," *Econometrica*, 37:424–438.

Granger, C. (1988): "Some Recent Developments in A Concept of Causality," *Journal of Econometrics*, 39:199–211.

Granger, C. (1986): "Developments in the Study of Cointegrated Economic Variables," *Oxford Bulletin of Economics and Statistics*, 48:213–228.

Granger, C. and Newbold, P. (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2:111–120.

Gweke, J. (1978): "Testing the Exogeneity Specification in the Complete Dynamic Simultaneous Equation Model," *Journal of Econometrics*, 7:163–185.

Gweke, J. (1982): "Measurement of Linear Dependence and Feedback Between Multiple Time Series," *Journal of the American Statistical Association*, 378:304–313.

Gweke, J., Meese, R., and Dent, W. (1983): "Comparing Alternative Tests of Causality in Temporal Systems," *Journal of Econometrics*, 21:161–194.

Hall, S. G. (1986): "An application of the Granger and Engle Two-Step Estimation Procedure To United Kingdom Aggregate Wage Data," *Oxford Bulletin of Economics and Statistics*, 48:229–239.

Harris, L. (1989): "S&P 500 Cash Stock Price Volatilities," *Journal of Finance*, 5:1155–1175.

Haugh, L. D. (1976, June): "Checking the Independence of Two Covariance-Stationary Time Series: A Univariate Residual Cross-Correlation Approach," *Journal of the American Statistical Association*, 71(354):378–385.

Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., and Lee, T. C. (1985): *The Theory and Practice of Econometrics*, New York: Wiley.

Koch, P. D., and Yang, S. S. (1986, June): "A Method for Testing the Independence of Two Time Series That Accounts for a Potential Pattern in the Cross-Correlation Function," *Journal of the American Statistical Association*, 81(394):533–544.

Kwaller, I. G., Koch, P., and Koch, T. (1987, December): "The Temporal Relationship Between S&P 500 Futures Prices and the S&P 500 Index," *Journal of Finance,* 42:1309–1389.

Mackinlay, A. C., and Ramaswamy, K. (1988, Summer): "Index Futures Arbitrage and the Behavior of Stock Index Futures Prices," *Review of Financial Studies,* 1:137–158.

MacKinnon, J. G. (1991): "Critical Values for Cointegration Tests," in *Long Run Economic Relationships: Readings in Cointegration,* Engle, R. F., and Granger, C. W. J., (eds.), Oxford University Press, pp. 267–276.

Milonas, N. T. (1986): "Price Variability and the Maturity Effect in Futures Markets," *The Journal of Futures Markets,* 6:443–460.

Modest, D. and Sundaresan, M. (1983, Summer): "The Relationship Between Spot and Futures Prices: Some Preliminary Evidence," *The Journal of Futures Markets,* 3:15–41.

Ng, N. (1987, May): "Detecting Spot Price Forecasts in Futures Prices Using Causality Tests," *The Review of Futures Markets,* 6:250–267.

Pierce, D. A. (1979): "R^2 Measures For Time Series," *Journal of the American Statistical Association,* 368:901–909.

Pierce, D. A., and Haugh, L. D. (1977): "Causality in Temporal Systems: Characterizations and Survey," *Journal of Econometrics,* 5:265–293.

Roll, R. (1984, September): "A Simple Implicit Measure of the Effective Bid–Ask Spread in an Efficient Market," *Journal of Finance,* 4:1127–1139.

Schwert, G. W. (1989, April): "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics,* 7(2):147–159.

Sims, C. A., and Dickey, D. (1972): "Money, Income and Causality," *American Economic Review,* 62:540–552.

Stoll, R. H., and Whaley, R. E. (1990, December): "The Dynamics of Stock Index and Stock Index Futures Returns," *Journal of Financial and Quantitative Analysis,* 25(3):441–468.