Cointegration and Error Correction Models: Intertemporal Causality between Index and Futures Prices

Asim Ghosh

INTRODUCTION

Stock market prices have been analyzed over the past 40 years in many ways to determine whether price changes are forecastable or not. These efforts have met with little success. Lately a new technique, called cointegration, has been developed which appears to hold some promise. Granger (1981) introduced the concept of cointegration where two variables may move together although they are nonstationary. The rationale behind the concept of cointegration is that there exists a long-run equilibrium relationship between the two variables. In the short-run they may deviate from each other but market forces, government intervention, etc. will bring them back together. Engle and Granger (1987) extended this concept and showed that cointegrated series have an error correction representation and conversely. With the error correction representation, a proportion of the disequilibrium in one period is expected to be corrected in the next period.

Anderson, Granger, and Hall (1990) model the term structure of U.S. Treasury bills as a cointegrated system and develop the error correction representation. Barnhart and Szakmary (1991) find spot and forward exchange rates are cointegrated and estimate the appropriate error correction model. Bessler and Covey (1991) investigate whether the daily futures market for live cattle is cointegrated with cash cattle market from August 21, 1985 through August 20, 1986. Their findings indicate marginal support for the cointegration hypothesis. Copeland (1991) examines and finds that the Deutschmark, Pound, French Franc, Swiss Franc, and Yen are not cointegrated against U.S. dollar using daily data from 1976 through 1989.

The purpose of this study is to investigate whether the index spot and futures price changes are predictable or not with a new econometric methodology that marries the short-run dynamic adjustment and long-run relationships between economic variables. Is there any causal relationship between spot and futures price changes? What is the direction of causality? The study finds that index spot and futures prices are integrated processes. Error correction models are developed and shown to be statistically significant in most cases and potentially useful for forecasting index spot and futures prices.

I am grateful to Mark J. Powers and two anonymous referees for their helpful comments and suggestions. I express appreciation to Henry Frank for editorial assistance. I am responsible for any remaining errors.

Asim Ghosh is an Assistant Professor in the Department of Finance at Rider College, Lawrenceville, New Jersey.

The Journal of Futures Markets, Vol. 13, No. 2, 193–198 (1993)

CCC 0270-7314/93/020193-06

DATA AND METHODOLOGY

Sandard and Poor's (S&P) 500 index spot and futures intra-day prices are obtained from the Chicago Mercantile Exchange (CME) for the time period from January, 1988 through December, 1988. The data are filtered by collecting the simultaneous spot and nearby futures prices approximately every 15 minutes on every Wednesday. Wednesday is chosen to avoid any confounding effects such as day-of-the-week. Daily closing prices of Commodity Research Bureau (CRB) index spot and near term delivery of futures are collected from the New York Futures Exchange (NYFE) covering the time period from June 12, 1986, through December 31, 1989. One thousand and one hundred observations from the S&P 500 index and 858 from the CRB index are employed for estimation. 30 observations from the S&P 500 index and 40 observations from the CRB index are used for forecasting. The logarithms of the price series are analyzed in this study.

Consider two nonstationary series x_t and y_t. They are called integrated of order one, denoted by $I(1)$, if they become stationary after first difference. It is generally true that an arbitrary linear combination, $y_t - \alpha x_t$, where α is a constant, is also nonstationary. However, if there exists a constant α such that this linear combination is $I(0)$ rather than $I(1)$, then x_t and y_t are said to be cointegrated and α is called the cointegrating parameter. This implies a steady-state relationship between the variables. When two series are cointegrated, Engle and Granger show that they have an error correction representation. A test for the presence of a unit root is conducted by performing the following augmented Dickey–Fuller (ADF) (1981) regression:

$$\Delta x_t = a_o + a_1 x_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta x_{t-i} + \varepsilon_t$$

where enough lagged differences are added to ensure that ε_t is white noise. Critical values are taken from MacKinnon (1990). To test for cointegration between the spot and nearby futures prices, the following regression analysis is performed:

$$S_t = b_o + b_1 F_t + u_t \tag{1}$$

where S_t denotes the log of current spot price and F_t is that of futures price. Equation (1) is considered a long-run relationship between spot and futures prices. A test for a unit root in the estimated residuals $\hat{u}_t$ determines the absence of presence of cointegration. The augmented Dickey–Fuller regression for estimated residuals $\hat{u}_t$ is:

$$\Delta \hat{u}_t = a\hat{u}_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta \hat{u}_{t-j} + v_t \tag{2}$$

where v_t is white noise. If the test rejects the null hypothesis of no cointegration then cointegration implies and is implied by an error correction representation and is expressed by the following system of equations:

$$\Delta y_t = \gamma_1 + \beta_1 u_{t-1} + \text{lagged}(\Delta x_t, \Delta y_t) + \varepsilon_{1t} \tag{3}$$

$$\Delta x_t = \gamma_2 + \beta_2 u_{t-1} + \text{lagged}(\Delta x_t, \Delta y_t) + \varepsilon_{2t} \tag{4}$$

where $u_t = S_t - b_o - b_1 F_t$ and $|\beta_1| + |\beta_2| \neq 0$. Equations (3) and (4) describe the short-run as well as long-run dynamics of the equilibrium relationship between the spot and futures prices. The error correction model is expected to provide better forecasts than a naive model provides. This is explained in the next paragraph. Parameters of eqs. (3) and (4) are estimated by OLS after running an OLS regression on eq. (1) to collect the residuals which are used then in eqs. (3) and (4). A large number of lagged differences

Table I
AUGMENTED DICKEY–FULLER TEST ON SPOT (S) AND FUTURES (F) PRICES

	S&P 500 Index		CRB Index		Critical Value (10%)
	S^{a}	F	S^{b}	F	
Levels	−1.99	−2.08	−2.01	−2.08	-2.57^{c}
Differences	−6.92	−6.14	−5.56	−5.49	−2.57

[a]One thousand and one hundred observations.
[b]Eight hundred and fifty-eight observations.
[c]Significance levels are from MacKinnon (1990).

of spot and futures prices are used in the initial estimation stage. Following the approach of Engle and Granger, this procedure is repeated several times eliminating insignificant coefficients at every stage until the final model is derived.

The existence of an error correction model implies some Granger causality between the series, which means that the error correction model can be used for forecasting. The forecasting performance of the error correction model is compared to that of a benchmark naive model by means of root mean squared error (RMSE). This naive method uses as a forecast the most recent information available concerning actual value. The forecasting equation is

$$F_{t+i} = x_t \text{ where}$$

$$\begin{aligned} F_{t+i} &= \text{forecast for period } t + i \\ t &= \text{present period} \\ i &= \text{number of periods ahead being forecast} \\ x_t &= \text{latest actual value (for period } t) \end{aligned}$$

THE EMPIRICAL EVIDENCE

Pseudo t values which determine whether each series is a nonstationary $I(1)$ process are presented in Table I. Augmented Dickey–Fuller unit root test is employed to test the levels and differences of the series. It is assumed that there is no time trend present in the data generation process. Tests for the presence of a unit root in the levels of each series fail to reject the null hypothesis. The pseudo t values of the difference series reject the null hypothesis of the presence of a unit root quite convincingly.

The results of the tests of cointegration are presented in Table II. Cointegration requires that the residuals are stationary. The null hypothesis of no cointegration is examined by testing the residuals from cointegration regression. The null hypothesis of no cointegration is rejected at the 10% level of significance for both systems. Thus, the evidence appears to indicate that the spot series is cointegrated with contemporaneous futures series for S&P 500 and CRB indices.

As Engle and Granger show, cointegration implies and is implied by an error correction representation. Cointegration describes the long-run equilibrium relationship between the variables. An error correction mechanism forces the short-run deviation from equilibrium in one period to move towards equilibrium in the next period. Results of the estimates of the error correction models are presented in Tables III and IV. Error correction models for the S&P 500 index and CRB index spot versus futures are estimated over 1100 and 858 observations, respectively. Error correction representations show that $|\beta_1| + |\beta_2| \neq 0$.

Table II
AUGMENTED DICKEY-FULLER TEST ON COINTEGRATION

Regressand	S&P 500 Futures[a]	CRB Futures[b]	10% Critical Value
S&P 500 Spot	−3.83	—	−3.05[c]
CRB Spot	—	−3.85	−3.05

[a]One thousand and one hundred observations.
[b]Eight hundred and fifty-eight observations.
[c]Significance levels are from MacKinnon (1990).

Table III
OLS ERROR CORRECTION MODEL FOR
S&P 500 INDEX SPOT AND FUTURE PRICES

	Model for ΔS_t		Model for ΔF_t	
Regressor	Coefficient	*S.E.*	Coefficient	*S.E.*
Constant	0.0001	0.0001	0.0001	0.0001
u_{t-1}	−0.1452	0.0637	0.0726	0.0590
DF_{t-1}	0.3270	0.0771	—	—
DS_{t-1}	−0.3284	0.0770	—	—

Table IV
OLS ERROR CORRECTION MODEL FOR CRB INDEX SPOT AND FUTURES

	Model for ΔS_t		Model for ΔF_t	
Regressor	Coefficient	S.E.	Coefficient	S.E.
Constant	.0001	0.0002	0.0001	0.0002
u_{t-1}	0.0252	0.0385	0.1251	0.0486
ΔF_{t-1}	—	—	0.3477	0.0841
ΔF_{t-2}	—	—	−0.2569	0.0862
ΔF_{t-3}	—	—	0.2066	0.0784
ΔF_{t-6}	0.1750	0.0627	—	—
ΔF_{t-7}	—	—	−0.1269	0.0711
ΔS_{t-1}	—	—	0.3475	0.0841
ΔS_{t-2}	—	—	0.2557	0.0862
ΔS_{t-3}	—	—	0.2064	0.0784
ΔS_{t-6}	−0.1753	0.0626	—	—
ΔS_{t-7}	—	—	0.1273	0.0711

Error correction terms are statistically significant for ΔS_t in the S&P 500 index cases and ΔF_t in the CRB index cases. This is consistent with the theory. If the coefficient of the error correction term is statistically insignificant, it implies that this variable does not adjust to departures from equilibrium. The constant in each system is statistically insignificant which confirms the assumption that the process is not generated by a linear trend.

Estimation of an error correction model implies the existence of causality between spot and futures price changes. The estimated error correction models are used to develop

30 one-step ahead forecasts for the S&P 500 index and 40 similar forecasts for the CRB index. These forecasts are then compared with naive univariate forecasts. Table V presents the summary statistics for these forecasts. Forecast standard deviations of the naive model are higher. The error correction models have relatively smaller forecast standard deviations, thus leading to smaller root mean squared errors. The error correction models outperform the naive model. Futures prices Granger-cause cash prices in the case of the S&P 500 index. The error correction models reduce the root mean squared error for ΔS_t and ΔF_t by about 14% and 24%, respectively, for the S&P 500 index over the naive model. The reduction in root mean squared error is more pronounced for the CRB index.

RESULTS AND CONCLUSIONS

This study investigates whether or not index spot and futures price changes are forecastable using cointegration. Intra-day data for the S&P 500 index and daily closing prices for the CRB index are used in this study.

The results indicate that both systems are cointegrated and therefore both systems exhibit a stable long-run equilibrium relationship. Resulting error correction models are statistically significant for most cases. Evidence appears to support the hypothesis that a proportion of the disequilibrium in one period is corrected in the next period.

For the S&P 500 index the adjustment of ΔS_t depends to a great extent on ΔF_t. The opposite is true for the CRB index. This evidence appears to suggest that more information flows from the futures to the spot market for the S&P 500 index and from the spot to futures for the CRB index.

The finding of cointegration between contemporary spot and futures prices in the S&P 500 and CRB markets provides evidence inconsistent with market efficiency.

Table V
SUMMARY STATISTICS FOR 1 STEP AHEAD FORECASTS

	S&P 500 Index			
	Naive		ECM.	
Statistic	ΔS_t	ΔF_t	ΔS_t	ΔF_t
Mean	−0.0004	−0.0004	0.0001	0.0001
SD	0.0012	0.0015	0.0003	$0.50 \cdot 10^{-4}$
MSE	$0.14 \cdot 10^{-2}$	$0.21 \cdot 10^{-2}$	$0.12 \cdot 10^{-2}$	$0.16 \cdot 10^{-2}$

	CRB Index			
	Naive		ECM.	
Statistic	ΔS_t	ΔF_t	ΔS_t	ΔF_t
Mean	0.0003	0.0003	0.0001	0.0001
SD	0.0041	0.0042	0.0002	0.0007
MSE	$0.62 \cdot 10^{-2}$	$0.67 \cdot 10^{-2}$	$0.41 \cdot 10^{-2}$	$0.41 \cdot 10^{-2}$

	RMSE Ratios			
	S&P 500		**CRB**	
Ratio	ΔS_t	ΔF_t	ΔS_t	ΔF_t
ECM/NAIVE	0.8571	0.7619	0.6613	0.6119

Cointegration between contemporaneous spot and futures implies that one price will help in forecasting the other. This is demonstrated in the article. Forecasts produced from the estimated error correction models are better than naive forecasts based on root mean squared error (RMSE) comparisons. Chowdhury (1991) finds evidence which seems to indicate the rejection of the "efficient market" hypothesis for four nonferrous metals—copper, lead, tin, and zinc—traded in the London Metal Exchange (LME).

Market efficiency, in the absence of transactions costs, implies that given current price, no additional information will improve the forecast of next period's price. If the S&P 500 futures contract trades in an efficient market the best forecast of next period's futures price is the price this period.

If the contemporaneous S&P 500 index and S&P 500 futures are cointegrated then the error correction model would suggest that the best forecast of next period's S&P 500 futures price is based on this period's S&P 500 futures price, the lagged residual error, and lagged changes in futures prices and index values.

Because transactions costs can not be ignored, it is not possible to reject market efficiency from the finding of cointegration between contemporaneous spot and futures prices. Market efficiency could be rejected if trading rules based on the error correction model provide positive profits after adjustment for the risk incurred in trading.

From the evidence presented, the error correction models outperform the naive model. It appears that this modeling strategy offers potential for forecasting price changes.

Bibliography

Anderson, H. M., Granger, C. W. J., and Hall, A. D. (1990): "Treasury Bill Yield Curves and Cointegration," Discussion Paper 90–24, University of California, San Diego.

Barnhart, S. W., and Szakmary, A. C. (1991): "Testing the Unbiased Forward Rate Hypothesis: Evidence on Unit Roots: Cointegration, and Stochastic Coefficients," *Journal of Financial and Quantitative Analysis,* 26:245–267.

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

Chowdhury, A. R. (1991): "Futures Market Efficiency: Evidence from Cointegration Tests," *The Journal of Futures Markets,* 11:577–589.

Copeland, L. S. (1991): "Cointegration Tests with Daily Exchange Rate Data," *Oxford Bulletin of Economics and Statistics,* 53:185–198.

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

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

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

Granger, C. W. J. (1981): "Some Properties of Time Series Data and Their Use in Econometric Model Specification," *Journal of Econometrics,* 16:121–130.

MacKinnon, J. G. (1990): "Critical Values for Cointegration Tests," Forthcoming in *Modelling Long-Run Economic Relationships,* Engle, R. F., and Granger, C. W. J. (ed.). Oxford University Press.