5-1998

An Examination of the Relationship Between Stock Index Cash and Futures Markets: A Cointegration Approach

Michael A. Pizzi
Ursinus College

Andrew J. Economopoulos
Ursinus College, aeconomopoulos@ursinus.edu

Heather M. O'Neill
Ursinus College, HONEILL@URSINUS.EDU

An Examination of the Relationship Between
Stock Index Cash and Futures Markets:
A Cointegration Approach

Michael A. Pizzi
Andrew J Economopoulos
Heather M. O'Neill

Last Revised: July 21, 1997

Michael A. Pizzi is the Ursinus College Galdfelter Research Scholar
Andrew J Economopoulos is an Associate Professor of Economics at Ursinus College
Heather M. O'Neill is an Associate Professor of Economics at Ursinus College

Section I: Introduction

The existence of price discovery, market efficiency and market stability associated with spot and futures markets continues as a prominent discussion among academics, practitioners and regulators.   Numerous papers examine the role of price discovery in the futures markets for various types of commodities and financial assets. Generally, the studies by Garbade and Silber (1983), Herbst, McCormack and West (1987), Kawaller, Koch and Koch (1987) and Schroeder and Goodwin (1991) indicate price discovery occurs more significantly in the futures market, as opposed to the cash market.

The literature develops over time using different econometric techniques, such as regression analysis and spectral analysis, to test efficiency and price discovery. Since the topics deal with short run and long run deviations from a presumed equilibrium relationship based on no arbitrage price bounds, the introduction of cointegration analysis with error correction models is fortuitous. The use of cointegration analysis and error correction models enable us to distinguish between short run deviations from equilibrium indicative of price discovery and long run deviations that account for efficiency and stability.

In this paper, we examine these issues - market efficiency[1], price discovery and market stability - using the intraday, minute by minute Standard and Poor's 500 (S&P 500) cash index and its 3-month and 6-month stock index futures. We undertake cointegration analysis and develop several error correction models.   The data extend over a 3-month contract's expiration period. Antoniou and Garrett (1993) use minute data and a cointegration model, but examine only the two days of data. Stoll and Whaley (1990) use five minute intervals for price announcements over a longer time period, but apply standard econometric analysis that fails to distinguish between short run and long run movements in indexes.   While Wahab and Lashgari (1993) introduced the use of cointegration techniques with stock price data, they used daily closing prices.   By using the finer grid of per minute data, we have a more robust test for cross-market efficiency.   We also incorporate nonsynchronous trading issues within our testing procedures similar to Stoll and Whaley (1990).   Unlike other papers, we model two different contract expirations, the 3 and 6-month contracts,

---

[1] Throughout this paper we use the term "market efficiency" to denote absences of the arbitrage opportunities between the spot and future stock index markets.

with the cash market to test whether these markets are efficient. Lastly, we analyze the data using cointegration techniques. Our paper is unique in that it incorporates the finer grid over an entire contract period using cointegration techniques.

Section II: Cointegration Analysis and Error Correction Models

To look for evidence of price changes in one market generating price changes in the other market so as to bring about a long run equilibrium relationship, we can write:

$$F_t - \beta_0 - \beta_1 S_t = e_t \qquad (1)$$

where S_t and F_t are contemporaneous cash and futures prices at time t, β_0 and β_1 are parameters and e_t is the deviation from parity. Ordinary least squares (OLS) is inappropriate if S_t and/or F_t are nonstationary because the standard errors are not consistent. The inconsistency disallows hypothesis testing of the cointegrating parameter β_1. According to Engle and Granger (1987) if S_t and F_t are nonstationary, which is suspected, but the deviations, e_t, are stationary, S_t and F_t are cointegrated. Thus, an equilibrium relationship exists between S_t and F_t. If the deviations are nonstationary, then either the markets are inefficient, in that an equilibrium relationship does not exist between them, or the two markets do not represent the same underlying asset.

For S_t and F_t to be cointegrated, they must be integrated of the same order. The order of integration is the number of times the series needs to be differenced in order to become stationary. Performing unit root tests on each univariate price series will determine the order of integration. If each series is nonstationary in the levels, but the first differences and the deviations, e_t, are stationary, the prices are cointegrated of order (1,1), denoted CI(1,1), with β_1 as the cointegrating coefficient.

An error correction model exists for each series, which is not subject to spurious results, if the two series are CI(1,1). Wahab and Lashgari (1993) state, "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" (p. 713). The first differences of the

2

cash and futures prices are ΔS_t and ΔF_t, respectively. Following Wahab and Lashgari (1993), these are then used in the error correction models of the form:

$$\Delta S_t = \alpha_1 + \alpha_S \hat{e}_{t-1} + \sum_{i=1}^{n} \alpha_{11}(i)\Delta S_{t-i} + \sum_{i=1}^{n} \alpha_{12}(i)\Delta F_{t-i} + \varepsilon_{St} \qquad (2)$$

$$\Delta F_t = \alpha_2 + \alpha_F \hat{e}_{t-1} + \sum_{i=1}^{n} \alpha_{21}(i)\Delta S_{t-i} + \sum_{i=1}^{n} \alpha_{22}(i)\Delta F_{t-i} + \varepsilon_{Ft} \qquad (3)$$

Equations (2) and (3) represent a near vector autoregression (VAR) in first differences, thus all variables are held jointly endogenous and OLS is an appropriate method of estimation. Each equation can be interpreted as having two parts.[2] The first part, $\hat{e}_{t-1}$, is the equilibrium error. This measures how the left hand side variable adjusts to previous period's deviation from long run equilibrium. The remaining portion of the equations are the lagged first differences, which represent short run effects of the previous period's changes in price on the current period's change in price.

The coefficients on the equilibrium errors, α_S and α_F, are the speed of adjustment coefficients. The speed of adjustment coefficients have important implications in an error correction model. At least one speed of adjustment coefficients must be non-zero in order for the model to be an error correction model. If the value of α_S in equation (2) is zero the current period change in the index does not respond at all to last period's deviation from long run equilibrium. If α_s is zero and all $\alpha_{12}(i)$ are zero then ΔF_t does not Granger cause ΔS_t. Wahab and Lashgari (1993) state two purposes for the speed of adjustment coefficients. They serve the role of identifying the direction of causal relation and show the speed at which departures from equilibrium are corrected.

Section III: Data and Econometric Testing

---

[2] In some error corrction models, the contemporaneous variable is included on the right-hand side of the equation making it a simultaneous sytem of equations. These models are generally employed in the macroeconomic literature where a two-stage least squares methodogoly is used to construct a predicted RHS endogenous variable. Since stock prices are assumed to follow a random walk, the construct of a proxy would be extremely difficult. The results of the estimation model that included such proxies would be called into question.

The Chicago Mercantile Exchange provided us with every price recorded in the S&P 500 stock index, as well as the transaction prices of the 3-month and 6-month S&P 500 index futures contract. The data are between January 1987 and March 1987. While the S&P 500 index is recalculated and transmitted to Chicago about every fifteen seconds, futures contracts prices may not change as often, especially for the 6-month expiration contract.

Given the non-uniform time periods in which price changes can occur, we calculated the mean prices for one minute intervals. The data begin after 8:40 AM (CST) and end at 3:00 PM (CST). Although the exchanges are open and record transactions both before and after our designated cut-offs, we do so to eliminate the stale price effects.

As Wahab and Lashgari (1993) point out, the lagged differences for the spot and futures prices, ΔS_t and ΔF_t, must be purged of serial correlation to eliminate the effects of infrequent trading and the bid/ask price effect. The methodology that follows is similar to Stoll and Whaley (1990).

Taking the log of each variable and its first difference, we represent the instantaneous relative price changes (returns) as:

$$s_t = \ln S_t - \ln S_{t-1} \qquad (4)$$

$$f_t = \ln F_t - \ln F_{t-1} \qquad (5)$$

Stoll and Whaley (1990) demonstrate that the effects of infrequent trading in the stock index can be modeled in terms of a pure autoregressive (AR) process and that the bid/ask price effect can be modeled in terms of a pure moving average (MA) process. The cash market, which is subject to infrequent trading, was purged of serial correlation with an AR(28). The three-month and six-month futures indexes, which potentially suffer from the bid/ask effects, required MA(25) and MA(30), respectively, to purge the effects.

Table 1 summarizes the serial correlation of the innovations in the transformed data. These innovation s_t' and f_t', replace ΔS_t and ΔF_t in the error correction model's equations (2) and (3).

4

TABLE 1

SUMMARY OF SERIAL CORRELATIONS OF INNOVATIONS
OF TRANSFORMED DATA
(January 2. 1987 to March 20. 1987)

VARIABLE (TRANSFORMED PROCESS) LAG(30)	LAG(6)	LJUNG-BOX LAG(12)	LAG(18)	LAG(24)	
SPOT INDEX	0.05	0.34	3.04	4.65	18.3
AR(28)	(1.00)	(1.00)	(1.00)	(1.00)	(0.95)
3M FUTURE	1.38	2.58	5.37	9.72	30.31
MA(25)	(0.97)	(1.00)	(1.00)	(1.00)	(0.45)
6M FUTURE	0.86	3.70	6.45	8.32	13.53
MA(30)	(0.99)	(0.99)	(0.99)	(1.00)	(1.00)

() Denotes the probability of accepting the null hypothesis of no serial correlation.

Note: The transformation process was applied to each day in order to avoid overnight effects. The test for serial correlation was applied to the residuals of the full series.

In order to determine the order of integration of each price series unit root tests were computed for each day on the levels of each price series. Three unit root tests were utilized; the Augmented Dickey-Fuller τ-test, the Phillips-Perron z-test, and the Weighted Symmetric τ-test. Performing all three tests on each day on the first differences of each series showed that the null hypothesis of a unit root was rejected for every day, thus we conclude each series is I(1).

Since we conclude that all three series are I(1), we test for cointegration with the following cointegrating regressions for the three month and six month futures, respectively.

$$F_{3t} = \beta_0 + \beta_1 S_t + e_{3t} \qquad (6)$$

$$F_{6t} = \beta_0 + \beta_1 S_t + e_{6t} \qquad (7)$$

According to Enders (1995), for large sample sizes it is only necessary to compute cointegrating equations in which either the spot index level or the futures level is

5

on the left hand side; asymptotic theory states that in large samples the position of the variables in the cointegrating equation does not matter.[3]

The Engle-Granger τ-test was performed on the $\{e_{3t}\}$ and $\{e_{6t}\}$ from equations (6) and (7). The results are reported in Table 2. We find that the spot price level and three month futures price level are CI(1,1) and the spot price level and six month futures price levels are CI(1,1).

Since both residual sequences are stationary, we estimate the following error correction models, using OLS regression, for the three month and six month futures, respectively. Table 3 displays the estimates of the speed of adjustment coefficients.

$$s'_t = \alpha_1 + \alpha_{3s_t}\hat{e}_{3t-1} + \sum_{i=1}^{30}\alpha_{11}(i)s'_{t-i} + \sum_{i=1}^{30}\alpha_{12}(i)f'_{3t-i} + \varepsilon_{s't} \qquad (8)$$

$$f'_{3t} = \alpha_1 + \alpha_{3f_t}\hat{e}_{3t-1} + \sum_{i=1}^{30}\alpha_{11}(i)s'_{t-i} + \sum_{i=1}^{30}\alpha_{12}(i)f'_{3t-i} + \varepsilon_{f't} \qquad (9)$$

$$s'_t = \alpha_1 + \alpha_{6s_t}\hat{e}_{6t-1} + \sum_{i=1}^{30}\alpha_{11}(i)s'_{t-i} + \sum_{i=1}^{30}\alpha_{12}(i)f'_{6t-i} + \varepsilon_{s't} \qquad (10)$$

$$f'_{6t} = \alpha_1 + \alpha_{6f_t}\hat{e}_{6t-1} + \sum_{i=1}^{30}\alpha_{11}(i)s'_{t-i} + \sum_{i=1}^{30}\alpha_{12}(i)f'_{6t-i} + \varepsilon_{f't} \qquad (11)$$

For the 3-month futures/cash index equations (8) and (9), the speed of adjustment coefficients indicate that the three month futures contract behaves somewhat differently than the six month futures contract. The significance of α_{3s_t} means that the spot market does respond to the previous period's deviation from equilibrium. A one standard deviation shock in the equilibrium error results in about a two percent change in the spot market innovation, indicating that the response is fairly large in

---

[3] The sample size was over 21,000 observations. The models, however, were tested using the both the spot and future as the left hand side variable. The results were identical. Only one set of results are reported.

6

TABLE 2
COINTEGRATION TESTS
BETWEEN SPOT, AND 3M & 6M FUTURES

DEPENDENT VARIABLE	COEFFICIENTS ON INDEPENDENT VARIABLES			E-G (tau) TEST (LAGS)
	SPOT	3M FUTURE	6M FUTURE	
WITH CONSTANT				
SPOT		-1.018		-6.79* (51)
3M FUTURE	-1.019			-6.81* (51)
SPOT			-0.993	-6.04* (58)
6M FUTURE	-0.995			-6.06* (58)
WITHOUT CONSTANT				
SPOT		-0.998		-5.40* (58)
3M FUTURE	-0.998			-5.40* (58)
SPOT			-0.992	-6.02* (58)
6M FUTURE	-0.992			-6.02* (58)

Cointegrating equations are bivariate models.

E-G denotes the Engle-Granger test of the residuals of the cointegration equation. The null hypothesis: Ho=unit root.
* denotes significance at the 1% level.

Table 3
Estimates of Coefficients

	Equation 8		Equation 9		Equation 10		Equation 11
$\alpha_{3s't}$	1.67E-7	$\alpha_{3f't}$	-9.53E-8	$\alpha_{6s't}$	1.64E-7	$\alpha_{6f't}$	-1.33E-7
t-statistic	6.69**		-1.43		6.85**		-2.17*
$\Sigma\,\alpha_{11}$	-1.10		0.612		-1.03		0.550
F-Statistic	195.97**		8.44**		180.90**		7.70**
$\Sigma\,\alpha_{12}$	0.643		-0.254		0.643		-0.178
F-Statistic	401.14**		8.76**		385.45**		4.46*

* denotes significance at the 5% level.
** denotes significance at the 1% level.

magnitude.[4] The lack of significance of α_{3ff} indicates that the current period three month futures innovation does not respond to the previous period's deviation from equilibrium. This means that any adjustment in the current period's futures innovation is caused by the lagged futures and cash market innovations.

Both speed of adjustment coefficients are significant in the error correction model using the six month futures innovations.[5] This means that both the current period spot and futures innovations respond to the previous period's deviation from equilibrium. Once again a one standard deviation shock in the equilibrium error results in approximately a two percent change in magnitude of either innovation.

The results of the error correction models do not support the theory that there is unidirectional causation from either market. The insignificant speed of adjustment coefficient in equation (9) does not mean that the spot market is not leading the futures market. All $\alpha_{11}(30)$ in equation (10) would have to be individually and jointly equal to zero to conclude that the spot market never leads the three month futures market. The F-statistic indicates that the we can reject the null hypothesis that the coefficients are jointly equal to zero. The first three lags of the index innovations $[\alpha_{11}(1),\ \alpha_{11}(2),\ \alpha_{11}(3)]$ are statistically significant in equation (10).[6] This means that the spot market leads the three month futures by at least 3 minutes. The last statistically significant index innovation occurs at lag 23 in equation (9). From this we conclude that the spot market leads the three month futures market by at least 3 minutes and at most 23 minutes.

Equation (8) demonstrates the leadership effect of the three month futures contract. The three month futures innovation shows a much stronger tendency to lead with the first twenty lagged futures market innovations being significant. The last

---

[4] The speed of adjustment coefficient size appears small because the error correction term is calculated as a residual from a regression on price levels, expressed with 5 digits, (e.g. an S&P500 price of 345 is 34500) and the innovations are residuals from an AR or MA model estimated on minute returns.

[5] The cross-maturity spread activities between the three month and six month, which are not directly modeled in this paper, may account for the significant coefficients in equation 10 & 11. This, however, is an area of future research.

[6] The full output from the estimation of the error correction models is available from the authors upon request.

statistically significant coefficient appears at lag 29. From this we conclude that the three month futures market leads the spot market by at least twenty minutes and at most by 29 minutes.

Turning to the six month futures contract we see that both markets are adjusting to long run equilibrium via the speed of adjustment coefficients. Equation (10) shows that six month futures innovation are significant to lag 20, with the last significant lag occurring at lag 29. This indicates that the six month futures contract tends to lead the spot market by at least 20 minutes and at most by 29 minutes. It is rather striking that both the three month futures and six month futures have the same leadership characteristics in relation to the spot market.

Equation (11) shows significant cash index innovations through lag 4 with last significant coefficient occurring at lag 18. From this we conclude that the spot market leads the six month futures market by at least 4 minutes and at most by 18 minutes. [7]

Section IV: Summary and Conclusion

In this paper we examined the relationship between the S&P 500 stock index and its respective futures contract. We examined both the three month and six month futures expiration over the same time period. Using several unit root tests we concluded that each price series was nonstationary in the levels but stationary after first differencing.

We tested both the spot index and three month futures and the spot index and six month futures for cointegration using the Engle-Granger two step procedure. We found that both the spot index and the three month futures and the spot and six month futures were cointegrated, indicating market efficiency. Thus, we calculated the two appropriate error correction models. The speed of adjustment coefficients indicated stability, but were smaller than expected.

---

[7] It should be noted that the residuals from equations (8-11) were examined via Yule-Walker methods for the presence of serial correlation. No significant serial correlation coefficients were found.

The results of these models showed that both the three and six month futures markets lead the spot market by at least 20 minutes. The spot market was found to lead the three month futures by at least 3 minutes and the six month futures by at least 4 minutes. While the futures market does tend to have a stronger lead effect, unidirectional causation of futures-to-spot is refuted.

References

Antoniou, A. and Garrett, I., (1993) "To What Extent Did Stock Index Futures Contribute to the October 1987 Stock Market Crash." *The Economic Journal,* (103):1444-1461

Enders, W. *Applied Econometric Time Series.* New York.: Wiley, 1995

Engle, R.F. and Granger C.W.J., (1987) "Cointegration and Error Correction: Representation, Estimation, and Testing." *Econometrica,* (55)2: 251-276

Garbade, K. D. and Silber, W. L., (1983) "Price Movements and Price Discovery in Futures and Cash Markets" *Review of Economics and Statistics,* (65)2: 289-297.

Herbst, A. F., McCormack, J. P. and West, E. N., (1987) "Investigation of a Lead-Lag Relationship Between Spot Stock Indices and Their Futures Contracts,"*Journal of Futures Markets,* (7)4: 373-381.

Kawaller, I. G., Koch, P. D. and Koch, T. W., (1987) "The Temporal Price Relationship between S&P500 Futures and the S&P500 Index," *The Journal of Finance,* (42)5: 1309-1329.

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

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

Wahab, M. and Lashgari, M., (1993) "Price Dynamics and Error Correction in Stock Index and Stock Index Futures Markets: A Cointegration Approach," *Journal of Futures Markets,* (13)4: 711-742.