Cointegration, Error Correction, and Price Discovery On Informationally Linked Security Markets

Article *in* Journal of Financial and Quantitative Analysis · December 1995

DOI: 10.2307/2331277 · Source: RePEc

4 authors:

Thomas H. Mcinish
The University of Memphis

Frederick H. deB. Harris
Wake Forest University

Gary L Shoesmith
Wake Forest University

Robert Wood
The University of Memphis

Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets

Frederick H. deB. Harris, Thomas H. McInish, Gary L. Shoesmith, and Robert A. Wood*

Abstract

Using synchronous transactions data for IBM from the New York, Pacific, and Midwest Stock Exchanges, we estimate an error correction model to investigate whether each of the exchanges is contributing to price discovery. Johansen's test yields two cointegrating vectors, which together verify the expected long-run equilibrium of equal prices across the three exchanges. Two error correction terms specified as the differences from IBM prices on the NYSE indicate that adjustments maintaining the long-run cointegration equilibrium take place on all three exchanges. That is, IBM prices on the NYSE adjust toward IBM prices on the Midwest and Pacific Exchanges, just as Midwest and Pacific prices adjust to the NYSE.

I. Introduction

Price discovery is the process by which markets attempt to find equilibrium prices (Schreiber and Schwartz (1986)).[1] In the U.S., stocks may trade on the New York, American, or one of the regional stock exchanges, or in third-market trading through the NASDAQ system. All the quotes and trades on these markets are informationally linked through the National Intermarket Trading System and the Consolidated Tape. This paper uses an error correction model to investigate the nature and extent to which regional exchanges contribute to the price discovery process.[2]

*Harris and Shoesmith, Babcock Graduate School of Management, Wake Forest University, Winston-Salem, NC 21709; McInish and Wood, Fogelman College of Business and Economics, Memphis State University, Memphis, TN 38152, respectively. This research was supported in part by the research grants program of the Babcock Graduate School of Management, Wake Forest University. The authors thank Clive Granger, Peter Bossaerts, Phil Dybvig, Robert Engle, Joel Hasbrouck, JFQA Managing Editor Jonathan Karpoff, and JFQA Referee Kenneth Kroner for their helpful comments and perceptive advice.

[1]Bronfman and Schwartz (1990) and Handa and Schwartz (1991) model the price discovery process, whereas Domowitz (1992) studies price discovery in automated trading systems.

[2]Hasbrouck (1991), (1993) has recently developed a different Bayesian VAR (BVAR) concept of price discovery—i.e., the proportion of the total innovation variance in the implicit efficient price accounted for by an exchange's own price variance. This approach produces interesting interpretations

Blume and Goldstein (1991) report that "most of the time the NYSE had the best quote." Lee (1993) discusses the possibility that regional exchanges "skim the cream," leaving less desirable trades for the NYSE. And Gardner and Subrahmanyam (1994) find evidence that harder, more informed trades are executed on the NYSE. McInish and Wood (1992a), (1992b) report, however, that enhanced competition from the regional exchanges lowers spreads and conclude that there is "clear evidence the regional exchanges are contributing to the price discovery process and are not merely free riding on primary exchange quotations" ((1992b), p. 17). In an earlier study, Garbade and Silber ((1979), p. 460) also found that "transactions prices on regional exchanges do contain information relevant for NYSE traders."

Recent advances in multivariate cointegration and error correction modeling provide a useful framework for analyzing equilibrium price adjustments in informationally linked markets. Johansen's (1988) multivariate test for cointegration is applied to three IBM stock price series based on synchronized trades on the New York, Pacific, and Midwest Stock Exchanges. All three exchanges employ specialist trading for hundreds of IBM trades per day. Two cointegrating vectors are found, which together identify the expected long-run equilibrium of equal prices across the three exchanges. An error correction model is then specified with two error correction terms, one computed as the price differential between the NYSE and the Pacific Exchange and the other as the price differential between the NYSE and the Midwest Exchange. The results show that error-correcting price adjustments take place on all three exchanges; NYSE prices adjust toward prices on the Midwest and Pacific Exchanges, just as Midwest and Pacific prices adjust to the NYSE.

The paper is organized as follows. Section II explains why cointegration and error correction are particularly relevant to price discovery. Section III describes the data collection procedure for matching trades to achieve synchronicity. Section IV presents the cointegration test results, followed by the estimation of the error correction model. A summary and conclusion are provided in Section V.

II. Cointegration, Error Correction, and Price Discovery

Engle and Granger (1987) (EG hereafter) provide the cornerstone research linking cointegrated series that move together through time to the concept of error correction—i.e., given movement away from long-run cointegration equilibrium in one period, a proportion of the disequilibrium is corrected in the next period. However, EG's two-step procedure has since been shown to be most appropriate for systems of only two variables with one possible cointegrating vector. In the more practical case of several variables, the cointegration test by Johansen (1988) is preferred, since it identifies the space spanned by the cointegrating vectors.

Following Granger's (1986) discussion of cointegration and error correction, let x_t be a vector of n component time series, each I(1) such that the one-differenced vector series is a zero mean, purely nondeterministic stationary process. The vector

of the lead-lag structure of permanent price changes across informationally linked security markets. It is less appropriate, however, for our purposes—i.e., to model the error correction dynamics that maintain cointegration equilibrium by removing disparities between the markets.

x_t is said to be cointegrated if there exists an $n \times r$ vector α such that $z_t = \alpha' x_t$ is stationary, i.e., I(0). By the Granger Representation Theorem (EG), if two or more series in x_t are cointegrated, there exists an error correction representation,

$$A(B)(1-B)x_t = -\gamma z_{t-1} + u_t, \tag{1}$$

where B is the backshift operator, γ is a matrix of coefficients ($n \times r$) of rank r, z_{t-1} is $r \times 1$ based on $r \leq n-1$ equilibrium error relationships $z_t = \alpha' x_t$, and u_t is a stationary multivariate disturbance. Thus, the error correction model in equation (1) is essentially a vector autoregression in differences with r lagged error correction terms ($-\gamma z_{t-1}$) included in each equation. For x_t having deterministic components, n constant terms are also included.

As an example of error correction relevant to this study, consider two stock-price series, PNY_t on the NYSE and PMW_t on the Midwest Exchange. From equation (1), the error correction model is

$$\begin{aligned}
\text{(2a)}\quad \Delta PNY_t &= -\gamma_1 z_{t-1} + m_1 + a_{11}\Delta PNY_{t-1} + a_{12}\Delta PNY_{t-2} \\
&\quad + \ldots + a_{1p}\Delta PNY_{t-p} + a_{21}\Delta PMW_{t-1} + a_{22}\Delta PMW_{t-2} \\
&\quad + \ldots + a_{2p}\Delta PMW_{t-p} + u_{1t}, \\
\text{(2b)}\quad \Delta PMW_t &= -\gamma_2 z_{t-1} + m_2 + b_{11}\Delta PNY_{t-1} + b_{12}\Delta PNY_{t-2} \\
&\quad + \ldots + b_{1p}\Delta PNY_{t-p} + b_{21}\Delta PMW_{t-1} + b_{22}\Delta PMW_{t-2} \\
&\quad + \ldots + b_{2p}\Delta PMW_{t-p} + u_{2t},
\end{aligned}$$

where $z_{t-1} = PNY_{t-1} - \alpha PMW_{t-1}$, m_1 and m_2 are constant terms, and p is the lag length. The $-\gamma_1 z_{t-1}$ and $-\gamma_2 z_{t-1}$ terms represent the error correcting adjustments that maintain the long-run equilibrium relationship $PNY_t = \alpha PMW_t$, where presumably $\alpha = 1$. For example, if $PNY_{t-1} > PMW_{t-1}$ (positive z_{t-1}), then $-\gamma_1 z_{t-1}$ should be negative, yielding reduced PNY_t, while $-\gamma_2 z_{t-1}$ is positive, yielding increased PMW_t. In this illustration, error-correcting adjustments must occur in either the NY price or the Midwest price or in both in order to maintain the long-run equilibrium between the two series. If $-\gamma_1$ is insignificant, PNY_t does not respond to differences from PMW_t, and PNY_t is considered exogenous within the system. Similarly, if $-\gamma_2$ is insignificant, PMW_t is exogenous.

As an empirical example, in EG's error correction model of consumption and income, income was found to be "weakly exogenous" in that the error correction term in the income equation proved to be statistically insignificant. Although consumption responds to deviations from income, income did not respond to deviations from consumption.

A. Cointegration, Error Correction, and Price Discovery

The possibility that one variable in a system of n cointegrated series is exogenous (independent) within the error correction process motivates the use of error correction models in evaluating price discovery. The cointegrating vectors define the long-run equilibrium, while the error correction dynamics characterize the price discovery process, i.e., "the process whereby markets attempt to find equilibrium" (Schreiber and Schwartz (1986)). Specifically, if prices on the Midwest and Pacific Exchanges responded to deviations from NYSE prices, but NYSE prices did not respond to deviations from regional exchange prices, that would be evidence that the price discovery process is focused in New York.

In this application, it should be emphasized that the error correction dynamics involve only cross-market information flows revealed by adjustments to price disparity across the three markets. They do not involve innovations in IBM prices due to information revelations (Hasbrouck (1993)), all of which are contained in u_t. These innovations are assumed to be independent of the error-correcting price adjustments, which maintain cointegration equilibrium. Although the R^2 for each of the error correction equations is therefore expected to be relatively small, it is the resolution of any remaining price differentials after coincident but often unequal price changes that reveals which markets contribute to the price discovery process.

For example, if IBM is initially trading at 50 on all three exchanges and information on a new computer product alters IBMMW, IBMPC, and IBMNY by +5/8, +3/4, and +5/8, respectively, the error correction terms reveal the reactions on each exchange to the Pacific price differential. If it is possible to execute the next trade at the observed disparity—i.e., by buying at 50 5/8ths and selling at 50 3/4—an arbitrage opportunity exists. There are several reasons, however, why an apparent arbitrage spread may not be sufficient to cover trading costs. First, the price disparity is observed only ex post, and our cointegration results show that such disparities do not persist. That is, market makers quickly adjust their quotes in an error correction fashion. Second, to realize arbitrage profit, the price adjustment must be sufficiently predictable so that the execution returns can, on average, cover trading costs.[3]

III. Data Collection Procedures

Since either private information or nonsynchronicity in the data can cause observed prices on linked markets to diverge, we have been careful to construct sample data sets that minimize the problem of nonsynchronicity. In 1990, IBM was the most heavily traded security on the New York and Pacific Stock Exchanges and the fifth most heavily traded security on the Midwest Exchange. All IBM trades on these three exchanges are acquired from the Institute for the Study of Security Markets (ISSM) data base, where the trades with condition codes C, R, N, and Z are excluded, and the trades are error filtered. We form six different data sets of matched trade tuples to address synchronicity and related issues.

The first data set is referred to as REPLACE ALL and is constructed as follows. Beginning at the start of each trading day, as soon as a trade is obtained for IBM on each of the three exchanges, the most recent trade for the other two exchanges is acquired to form the first matched trade tuple. That tuple is then saved,

[3]Equality of the three prices is not assured. Cointegrated price series may diverge but remain arbitrage-free if the differentials next period are not predictable. For example, in the market microstructure models of Glosten and Milgrom (1985), Kyle (1989), and Back (1992), (1993), the information content of prices is endogenous and less than the pooled information of all agents. Brenner and Kroner (1995) contrast theoretical differences between arbitrage and cointegration equilibrium.

and a new matched trade tuple is formed in the same manner.[4] Consequently, each observation reflects new trades in all three markets and may reflect zero, one, two, or three new prices. Often no price will have changed from one tuple to the next.

Time in such a model is trading time rather than continuous clock time, and the frequency of the data is determined by the market with the fewest trades. IBM trades approximately 10 times in New York (243,000 per year) for each trade in the Midwest (24,000 per year). Pacific's trading frequency is intermediate, at 71,000 per year. This REPLACE ALL data collection procedure generated 16,369 matched trade tuples—an average of 80 observations per day.

This high frequency of microstructure data is crucial to testing for pricing dynamics across informationally linked security markets for two reasons. First, cointegration models capture "long-run" equilibrium relationships wherein time series can diverge temporarily but then readjust to persistent cointegrated patterns. One year of IBM trades is long run in the sense that more than 16,000 such price adjustments can occur. Also, our results in the next section show that these REPLACE ALL data are sufficient to obtain the expected long-run cointegrating equilibrium of price equality across the three exchanges. Second, we must guard against observation intervals so long that error correction takes place within rather than between the tuples. Just as annual data on household consumption and income cannot detect an error correction process reflecting monthly household budgets, so too, daily stock price data cannot detect the error correction from higher frequency trading strategies.

The reaction of the NYSE, Pacific, and Midwest Exchanges to price differentials can be detected accurately only when the matched trades observed are synchronous across the three markets. Therefore, as recommended by Blume and Goldstein (1993), we preadjust all trades for minimum reporting lags between New York time stamps and those on the regional exchanges (16 seconds on the Midwest, 5 seconds on the Pacific) before matching them in observation tuples.[5] A measure of synchronicity in such data sets is the seconds elapsed between the first recorded price and the last recorded price in the tuple (i.e., the SPAN). For the REPLACE ALL data set, the distribution of SPAN for IBM tuples in 1990 is positively skewed, with a modal value of 20 seconds, a mean value of 102 seconds, and a median value of 88 seconds. This means that once Midwest trades, the REPLACE ALL data collection procedure acquires New York and Pacific trading prices over the previous one to two minutes.

Two alternative data sets that reduced SPAN still further were also constructed. First, we attempted to minimize the observation interval by looking forward in trading time as well as backward.[6] Exhibit 1 illustrates the modifications in the REPLACE ALL data collection procedure. Once the Midwest market trades third at a_m, forming the first tuple by acquiring the most recent trades in the other two markets (i.e., a_p and a_n) fails to minimize the SPAN that would be available from

---

[4] That is, we first identify the last exchange to record an opening trade. To form the first tuple, we then match that trade with the most recent trade in each of the other two markets. For the second tuple, we again wait for all three markets to trade and then retrieve the most recent trade in the other two.

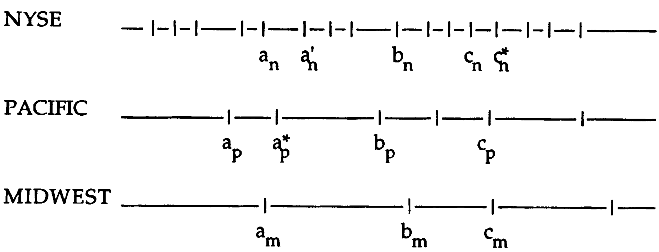
[5] As a check, we also ran results for the raw, unadjusted time stamp data.

[6] We are indebted to the referee for suggesting this minimal span approach to obtaining more synchronous data tuples.

looking forward to the next Pacific trade at a_p^*. That is, the REPLACE ALL tuple (a_p, a_n, a_m) has a longer SPAN than the alternative first tuple (a_n, a_m, a_p^*). This data collection procedure of minimizing the tuple span, which we refer to as MINSPAN, has a natural stopping rule associated with the third market to trade. After acquiring (a_p, a_n, a_m), we check forward to see whether a subsequent trade in either the first or second market to trade reduces the SPAN. If so, the tuple is updated as (a_n, a_m, a_p^*) and saved. If not, the original tuple is saved [e.g., b_p, b_n, b_m)], three new trades are acquired, and the SPAN minimizing inquiry is repeated. Hence, (c_p, c_m, c_n^*) updates (c_n, c_p, c_m). With such an approach, synchronicity is substantially improved; the MINSPAN procedure generated 13,154 observations with a mean SPAN of only 52.8 seconds.

EXHIBIT 1

Time Lines of Trades in Three Informationally Linked Security Markets*



Tuple	REPLACE ALL	MINSPAN	REPLACE OLDEST
1	(a_p, a_n, a_m)	(a_n, a_m, a_p^*)	(a_p, a_n, a_m)
2	(b_p, b_n, b_m)	(b_p, b_n, b_m)	(a_n, a_m, a_p^*)
3	(c_n, c_p, c_m)	(c_p, c_m, c_n^*)	(a_m, a_p^*, a_n')
4			(b_p, b_n, b_m)

*Data tuples are listed by time of trade from oldest to most recent under three different data collection procedures.

We also experimented with shortening the observation interval by a third data collection procedure. The REPLACE OLDEST procedure forms new tuples by replacing only the oldest trade in the preceding tuple. Exhibit 1 illustrates the additional data tuples this procedure generates between the two adjacent (a_p, a_n, a_m) and (b_p, b_n, b_m) REPLACE ALL tuples. As expected, the mean SPAN falls from 102 to 88 seconds. All three procedures for matching IBM trades form reasonably synchronous time series of IBMNY, IBMPC, and IBMMW prices for the NYSE, Pacific, and Midwest Exchanges. As the cointegration and error correction results for this third REPLACE OLDEST data set proved to be qualitatively identical,[7] we proceed with reporting results for the REPLACE ALL and MINSPAN data sets.

[7]The unreported REPLACE OLDEST results are available from the authors.

Because of its higher trading frequency, NYSE often appears in the middle of the tuple observation intervals. Table 1 shows the relative frequency with which each exchange appears first in the REPLACE ALL and MINSPAN tuples. For REPLACE ALL, Pacific is usually first (45.1 percent), followed by New York (28.7 percent), and then Midwest (26.2 percent). One spurious effect that could mask the price discovery process we are trying to detect is a systematic pattern of bid-ask bounce across the three markets. A priori, whether the next IBM trade anywhere will be at the bid or at the ask must be equally likely. However, if our data collection procedures impose a pattern of one exchange observed first, matched with another exchange observed second, and if the specialists bounce from a trade at the bid on one exchange to a trade at the ask on the next, then bid-ask bounce across the markets could be mistaken for error correction. To address this potential bias, we recollected three subsets of the REPLACE ALL data: tuples with New York first, Midwest first, and Pacific first.[8] The mean SPANs for all these tuples (see Table 1) were also quite low: 61, 104, and 107 seconds.

TABLE 1

Data Collection Procedures: Observation Intervals and Ordering of Tuples

	Span (seconds)	Frequency (first)
Randomly ordered tuples adjusted for minimum reporting lags		
REPLACE ALL (16,369 obs.)	101.86	
Midwest		26.2%
Pacific		45.1%
NYSE		28.7%
MINSPAN (13,154 obs.)	58.20	
Midwest		32.1%
Pacific		32.2%
NYSE		35.7%
Ordered tuples adjusted for minimum reporting lags		
MIDWEST FIRST (12,147 obs.)	104.23	100.0%
PACIFIC FIRST (11,384 obs.)	107.06	100.0%
NYSE FIRST (7,029 obs.)	61.02	100.0%

IV. Cointegration and Error Correction Results

The first step in testing for cointegration is to determine the order of integration of each series. The most common approach is the augmented Dickey-Fuller (ADF) test based on Dickey and Fuller (1979), (1981), which runs the regression,

$$(3) \quad \Delta x_{it} \;=\; -\delta x_{it-1} + \beta_0 + \beta_1 \Delta x_{it-1} + \beta_2 \Delta x_{it-2} + \ldots + \beta_j \Delta x_{it-j} + e_t.$$

If $\delta = 0$, x_{it} is I(1). If δ is positive and significant, x_{it} is I(0). The t-statistic reported for the coefficient $-\delta$ is used as the test statistic and compared with critical values derived from simulation (see Dickey and Fuller (1981)).

[8]To do so, we imposed a constraint on the REPLACE ALL procedure that, for example, New York be recorded first—i.e., that no New York trade could intervene between Midwest and Pacific trades.

Applying equation (3) to the three IBM stock price series (in natural log form) generated by the REPLACE ALL procedure with $j = 6$ initially and dropping all insignificant lags at the 0.05 level, the test fails to reject I(1) for x_{it} in each case. Regressing second differences ($\Delta^2 x_{it}$) on lagged Δx_{it} and six lagged $\Delta^2 x_{it}$, and dropping all insignificant lags, the t-statistics ($-\delta$s) on lagged ΔIBMMW, ΔIBMPC, and ΔIBMNY are −96.20, −97.14, and −131.70. Each is significant at the 1-percent level, rejecting I(1) for the three Δx_{it}. Thus, each IBM price series is I(1). Similar ADF test statistics were obtained for each of the other five data sets.

We determined the optimal system lag length for each of the three-equation models using a likelihood ratio procedure based on a chi-squared statistic for the unconstrained vector autoregression (VAR) model in levels. This procedure involves successively testing shorter lag lengths as restrictions against longer lag lengths, where the lag order is set uniformly across the three equations. Beginning with a lag length of $p = 6$ and moving to shorter lengths, the chi-squared statistic becomes significant with the elimination of the third lag using the REPLACE ALL data and the second lag with the MINSPAN data. Thus, the system lag length is $p = 3$ using the REPLACE ALL data and $p = 2$ with the MINSPAN data.

A. Cointegration Results

Table 2 shows the Johansen test results for the three IBM series obtained with the REPLACE ALL and MINSPAN data sets, including eigenvalues, eigenvectors, and test statistics for the maximal eigenvalue test. For both data sets, the null hypotheses of $r = 0$ and $r \leq 1$ cointegrating vectors are successively rejected by the maximal eigenvalue test at the 1-percent level, whereas $r \leq 2$ is not rejected. Thus, $r = 2$ cointegrating vectors is concluded under both data collection procedures. In each case, the two eigenvectors corresponding to the first two eigenvalues are the cointegrating vectors representing the empirical long-run relationships between IBMMW, IBMPC, and IBMNY.

The two equilibrium error relationships implied by the REPLACE ALL cointegrating vectors are

$$\text{(4a)} \qquad z_{1t} = 170.69\ \text{IBMMW}_t - 151.16\ \text{IBMPC}_t - 19.528\ \text{IBMNY}_t,$$

$$\text{(4b)} \qquad z_{2t} = -43.81\ \text{IBMMW}_t - 86.286\ \text{IBMPC}_t + 130.10\ \text{IBMNY}_t.$$

The sample means for z_{1t} and z_{2t} are -2.91×10^{-12} and 9.62×10^{-14}, with t-statistics for $H_0 : \mu_{z_1} = 0$ and $H_0 : \mu_{z_2} = 0$ of -2.18×10^{-9} and 7.00×10^{-11}. Setting $z_1 = 0$ and $z_2 = 0$ and substituting, the expected long-run equilibrium of IBMMW = IBMPC = IBMNY is obtained to three decimal places. The MINSPAN cointegration results yield the same equilibrium. To obtain this equilibrium, two cointegrating vectors are required, each having parameters that sum to zero. Any other possibility would allow for one or more series pairs to not be equal in the long run.

TABLE 2

Johansen Cointegration Test Statistics

Panel A. REPLACE ALL Data Collection Procedure

		Eigenvalues ($\hat{\lambda}$)	
	(0.174	0.163	0.000001)
		Eigenvectors ($\hat{V}$)	
IBMMW	170.688	−43.810	−0.333
IBMPC	−151.160	−86.286	−0.523
IBMNY	−19.528	130.098	−0.144
		Cointegration Test Statistics* (T = 16,365)	
	H_0	$\hat{\lambda}_{max}$	λ_{max} (0.99)
	$r \leq 2$	0.02	11.58
	$r \leq 1$	2,917.8	18.78
	$r = 0$	3,120.6	26.15

Panel B. MINSPAN Data Collection Procedure

		Eigenvalues ($\hat{\lambda}$)	
	(0.285	0.243	0.00003)
		Eigenvectors ($\hat{V}$)	
IBMMW	153.471	−18.354	−0.452
IBMPC	−127.307	−74.566	−0.396
IBMNY	−26.166	92.922	−0.152
		Cointegration Test Statistics* (T = 13,151)	
	H_0	$\hat{\lambda}_{max}$	λ_{max} (0.99)
	$r \leq 2$	0.3	11.58
	$r \leq 1$	3,658.4	18.78
	$r = 0$	4,407.2	26.15

*Critical values are taken from Johansen and Juselius (1990).

The maximal eigenvalue test statistics shown above indicate two cointegrating vectors for each of the REPLACE ALL and MINSPAN data sets. The data sets were constructed as illustrated in Exhibit 1. The optimal system lag length for each model was obtained using a likelihood ratio procedure based on a chi-squared statistic for the unconstrained vector autoregression (VAR) model in levels. This procedure resulted in a system lag length of $p = 3$ for the REPLACE ALL data and $p = 2$ for the MINSPAN data. For each data set, the two cointegrating vectors are given by the first two eigenvectors, the parameters of which sum to approximately zero. Combining the two equilibrium error relationships implied by the cointegrating vectors yields the expected long-run equilibrium relationship of IBMMW = IBMPC = IBMNY.

B. Error Correction Model Specification and Estimation

By equation (1), the cointegration results in Table 2 can be used to specify an error correction model to provide clear evidence on whether NYSE prices adjust to deviations from prices on the regional exchanges. Given that the long-run equilibrium obtained by each data set's cointegration results is IBMMW = IBMPC = IBMNY, each error correction model is estimated using the two price differentials involving the NYSE as the error correction terms, i.e., $Z1_t = \text{IBMNY}_t - \text{IBMPC}_t$

and $Z2_t$ = IBMNY$_t$ − IBMMW$_t$.[9] These error correction terms provide clear insight into the significance and magnitudes of the actual error correction pricing dynamics.

Beginning with the REPLACE ALL data set, the top half of Table 3 reports the estimation results for a VAR in first differences. The VAR in differences is equivalent to the error correction model, but without the two error correction terms. This comparison is included to observe the marginal effects of adding the error correction terms.[10] All estimation results are obtained using seemingly unrelated regression. Looking at the VAR results, all of the 27 coefficient estimates on lagged ΔIBMMW, ΔIBMPC, and ΔIBMNY are significant at the 1.0-percent level. These cross-price effects provide the first indication that independent information contained in the prices of each exchange influences prices on the other exchanges.

The estimated error correction model in the bottom half of Table 3 yields a number of important observations. First, each of the coefficient estimates on the lagged $Z1_{t-1}$ and $Z2_{t-1}$ terms is significant at the 1-percent level. Second, with the error correction terms included, 21 of the 27 lagged ΔIBMMW, ΔIBMPC, and ΔIBMNY terms are now insignificant, with the only significant lags being first-order lags. Third, the F-statistics show that the lagged equilibrium errors ($Z1_{t-1}$ and $Z2_{t-1}$) are responsible for most of the variation explained in ΔIBMMW, ΔIBMPC, and ΔIBMNY. This is also clear in comparing the R^2 statistics from the VAR in differences and the error correction model. With the addition of the error correction terms, explained variation increases by 83 percent, 55 percent, and 30 percent in the ΔIBMMW, ΔIBMPC, and ΔIBMNY equations. These findings provide unambiguous evidence that i) error-correcting price adjustments occur on all three exchanges in maintaining cross-market equilibrium, and ii) these adjustments account for a large share of the explained variation in the estimated ΔIBMMW, ΔIBMPC, and ΔIBMNY equations.

Looking at the magnitudes of the coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ in each equation provides several specific insights into the error correction process. First, the coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ in the ΔIBMNY equation are smaller in absolute value than the corresponding coefficients in the other two equations, suggesting that reactions on the NYSE to the price differentials defined by $Z1$ and $Z2$ are generally smaller than the reactions on the regional exchanges. This is especially obvious when comparing the coefficients on $Z2_{t-1}$ in the ΔIBMNY and ΔIBMMW equations (−0.256 vs. 0.433) and $Z1_{t-1}$ in the ΔIBMNY and ΔIBMPC equations (−0.201 vs. 0.475). Thus, ignoring cross effects for the moment, if IBMNY is 1 percent higher than IBMMW, error-correcting adjustments take place such that IBMNY declines by 0.256 percent while IBMMW increases by 0.433 percent. Likewise, given a 1-percent difference between IBMNY and IBMPC, IBMNY declines by 0.201 percent while IBMPC increases by 0.475 percent. Statistical tests confirm that the coefficients on $Z2_{t-1}$ in the ΔIBMNY

[9]The use of these error correction terms, rather than those defined by equations (4a) and (4b), was suggested by the referee. An error correction model based on (4a) and (4b) was also estimated. The coefficients on z_{1t} and z_{2t} were the correct sign and significant at the 1-percent level in each equation.

[10]Notably, by the Granger Representation Theorem, VAR in differences is a misspecification in this case, given the cointegration results in Table 2.

TABLE 3

Estimation Results: REPLACE ALL Data Set[a]

	Dependent Variable		
	ΔIBMMW	ΔIBMPC	ΔIBMNY
Panel A. VAR First Difference Model			
ΔIBMMW($t-1$)	−0.211(−19.38*)	0.261(23.49*)	0.258(18.21*)
ΔIBMMW($t-2$)	−0.213(−18.76*)	0.122(10.58*)	0.118(8.00*)
ΔIBMMW($t-3$)	−0.124(−11.49*)	0.091(8.22*)	0.060(4.28*)
ΔIBMPC($t-1$)	0.171(16.45*)	−0.333(−31.43*)	0.194(14.36*)
ΔIBMPC($t-2$)	0.124(11.31*)	−0.242(−21.66*)	0.115(8.07*)
ΔIBMPC($t-3$)	0.051(4.98*)	−0.153(−14.53*)	0.038(2.87*)
ΔIBMNY($t-1$)	0.096(12.87*)	0.122(16.15*)	−0.482(−49.86*)
ΔIBMNY($t-2$)	0.111(13.57*)	0.131(15.69*)	−0.267(−25.12*)
ΔIBMNY($t-3$)	0.067(9.00*)	0.066(8.68*)	−0.126(−12.93*)
CONSTANT	−0.90E-05(−0.87)	−0.91E-05(−0.85)	−0.11E-04(−0.80)
F-Stats: ΔIBMMW	186.97*	188.42*	111.99*
ΔIBMPC	99.93*	370.05*	72.08*
ΔIBMNY	81.08*	114.77*	829.53*
R^2	0.046	0.080	0.133
DW	2.019	2.022	2.019
Panel B. Error Correction Model			
ΔIBMMW($t-1$)	0.089(5.58*)	0.064(3.93*)	0.090(4.33*)
ΔIBMMW($t-2$)	0.73E-04(0.005)	−0.018(−1.29)	−0.0002(−0.009)
ΔIBMMW($t-3$)	0.003(0.27)	0.003(0.27)	−0.007(−0.47)
ΔIBMPC($t-1$)	0.002(0.16)	−0.0004(−0.03)	0.058(2.88*)
ΔIBMPC($t-2$)	0.002(0.13)	−0.011(−0.77)	0.027(1.52)
ΔIBMPC($t-3$)	−0.021(−1.93)	−0.022(−1.97)	−0.006(−0.43)
ΔIBMNY($t-1$)	−0.048(−4.13*)	−0.030(−2.50)	−0.146(−9.57*)
ΔIBMNY($t-2$)	0.010(0.99)	0.026(2.46)	−0.033(−2.51)
ΔIBMNY($t-3$)	0.011(1.30)	0.010(1.17)	0.003(0.30)
CONSTANT	−0.93E-05(−0.91)	−0.95E-05(−0.92)	−0.10E-04(−0.76)
$Z1(t-1)$ (NY-PC)	−0.235(−14.01*)	0.475(27.87*)	−0.201(−9.23*)
$Z2(t-1)$ (NY-MW)	0.433(25.34*)	−0.272(−15.65*)	−0.256(−11.52*)
F-Stats: ΔIBMMW	17.67*	12.92*	11.39*
ΔIBMPC	1.95	1.65	4.10*
ΔIBMNY	17.37*	14.65*	47.64*
$Z1$, $Z2$	341.92*	411.23*	397.23*
R^2	0.084	0.124	0.173
DW	2.000	2.000	2.000

[a]Estimates(t-statistics) are indicated for each variable.

*Indicates significance at the 1.0-percent level.

This table presents the error correction modeling results for the REPLACE ALL data set, along with similar results for a VAR in first differences, which is identical to the error correction model except that the error correction terms ($Z1$ and $Z2$) are excluded. Based on the equilibrium relationship IBMMW = IBMPC = IBMNY obtained from the cointegration test results in Table 2, $Z1$ and $Z2$ are defined as the two price differentials involving the NYSE, i.e., $Z1_t = \text{IBMNY}_t - \text{IBMPC}_t$ and $Z2_t = \text{IBMNY}_t - \text{IBMMW}_t$. Seemingly unrelated regression is used in estimating both models. The optimal system lag length is $p = 3$ (see Table 2). The error correction results show that each of the coefficient estimates on $Z1_{t-1}$ and $Z2_{t-1}$ is significant at the 1-percent level. This implies that price adjustments take place on all three markets in maintaining price equality. Also, the F-statistics show that $Z1_{t-1}$ and $Z2_{t-1}$ are responsible for most of the variation explained in ΔIBMMW, ΔIBMPC, and ΔIBMNY, which is also clear in comparing the R^2 statistics from the two estimated models. Finally, the magnitudes of the coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ provide a detailed account of the error correction process.

and ΔIBMMW equations are significantly different (in absolute terms) at the 1-percent level (t = 5.01), as are the coefficients on $Z1_{t-1}$in the ΔIBMNY and ΔIBMPC equations (t = 7.91).

Examining the dynamics involving both $Z1_{t-1}$ and $Z2_{t-1}$ in each equation, we find that the error correction terms appear to serve balancing roles in the price adjustment process. For example, suppose in period $t-1$ IBMMW = IBMPC, but IBMNY is 1 percent higher than both IBMMW and IBMPC ($Z1_{t-1} = Z2_{t-1} = 1$). The price adjustments in period t implied by the error correction terms are as follows. By the first equation, IBMMW increases a net 0.198 percent; 0.433 percent due to $Z2_{t-1}$ plus -0.235 percent due to $Z1_{t-1}$. Thus, the fact that both IBMMW and IBMPC are lower than IBMNY yields a smaller upward price change in IBMMW than if IBMPC = IBMNY. Similar dynamics yield a net increase of 0.203 percent for IBMPC in period t, virtually the same as IBMMW. Meanwhile, IBMNY declines by a larger 0.457 percent ($-0.201 - 0.256$), given its positive differential with both IBMMW and IBMPC. In all, within one period, the original 1-percent price differential is reduced to 0.345 percent for IBMNY vs. IBMMW and 0.34 percent for IBMNY vs. IBMPC.

As a second example, suppose IBMNY is in the middle in period $t-1$, 1 percent higher than IBMMW ($Z2_{t-1} = 1$), but 1 percent lower than IBMPC ($Z1_{t-1} = -1$). In this case, IBMMW increases by 0.668 percent $[-1(-0.235) + 0.433]$ in period t, while IBMPC declines by 0.747 percent $[-1(0.475) - 0.272]$. Meanwhile, IBMNY changes by only-0.055 percent ($0.201 - 0.256$). Thus, IBMNY changes very little, while the other two exchanges react quickly to close the gap. In this regard, tests show that the coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ in the ΔIBMNY equation are not significantly different (t = 1.34). Thus, price differentials with the Midwest and Pacific Exchanges appear to have roughly equal influence on IBMNY. For the ΔIBMMW and ΔIBMPC equations, the coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ (in absolute terms) are significantly different at the 1-percent level, which simply indicates that, in each case, the own-price differential with IBMNY is more influential than the cross-price differential with the other regional exchange.

These two examples give an explicit account of the nature and magnitudes of the error-correcting adjustments that maintain equilibrium IBM prices across the three exchanges. In summary, the information reflected in IBM prices on each of the three exchanges makes an independent and significant impact on IBM prices on the other two exchanges. The NYSE's reaction to price differentials with the regional exchanges is smaller, on average, than the ensuing reaction of the regional exchanges. As demonstrated by the first example, however, IBMNY may at times react more than either IBMMW or IBMPC, depending on IBMNY's original position within the pricing disequilibrium.

Table 4 shows the VAR and error correction modeling results for the MINSPAN data set, which are very similar to those in Table 3 for the REPLACE ALL data. One notable difference is that the coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ in the ΔIBMNY equation are significantly different in this case (t = 8.99), with IBMNY reacting more strongly to price differentials with IBMMW. Otherwise, the error correction results and implied pricing dynamics are quite similar. Repeating the first example above, the effects from the error correction terms on ΔIBMMW, ΔIBMPC, and ΔIBMNY for the MINSPAN data set are 0.173 percent, 0.145 percent, and −0.527 percent, which compares with 0.198 percent, 0.203 percent, and −0.457 percent for the REPLACE ALL data. For the second example with the MINSPAN data set, the effects on ΔIBMMW, ΔIBMPC, and ΔIBMNY are 0.703 percent, −0.959 percent, and −0.167 percent, which compares with 0.668 percent, −0.747 percent, and −0.055 percent for the REPLACE ALL data.

TABLE 4

Estimation Results: MINSPAN Data Set[a]

	Dependent Variable		
	ΔIBMMW	ΔIBMPC	ΔIBMNY
Panel A. VAR First Difference Model			
ΔIBMMW($t-1$)	−0.367(−25.39*)	0.264(17.77*)	0.208(11.66*)
ΔIBMMW($t-2$)	−0.166(−11.56*)	0.135(9.15*)	0.107(6.07*)
ΔIBMPC($t-1$)	0.226(16.34*)	−0.409(−28.86*)	0.190(11.20*)
ΔIBMPC($t-2$)	0.106(7.71*)	−0.177(−12.53*)	0.078(4.60*)
ΔIBMNY($t-1$)	0.148(15.64*)	0.146(15.07*)	−0.443(−38.01*)
ΔIBMNY($t-2$)	0.069(7.27*)	0.060(6.12*)	−0.203(−17.26*)
CONSTANT	−0.18E-04(−1.06)	−0.18E-04(−1.03)	−0.19E-04(−0.93)
F-Stats: ΔIBMMW	322.15*	158.89*	68.46*
ΔIBMPC	113.48*	416.64*	62.83*
ΔIBMNY	122.29*	113.75*	721.83*
R^2	0.054	0.066	0.099
DW	2.018	2.018	2.032
Panel B. Error Correction Model			
ΔIBMMW($t-1$)	−0.080(3.61*)	−0.012(−0.53*)	−0.010(−0.38)
ΔIBMMW($t-2$)	−0.020(1.21)	−0.006(−0.36)	−0.002(−0.11)
ΔIBMPC($t-1$)	0.049(2.32)	−0.045(−2.10)	0.075(2.97*)
ΔIBMPC($t-2$)	0.017(1.10)	0.005(0.34)	0.020(1.05)
ΔIBMNY($t-1$)	0.033(2.42)	0.051(3.61*)	−0.093(−5.54*)
ΔIBMNY($t-2$)	0.009(0.82)	0.011(1.02)	−0.020(−1.52)
CONSTANT	−0.18E-04(−1.08)	−0.18E-04(−1.06)	−0.19E-04(−0.93)
$Z1(t-1)$ (NY-PC)	−0.265(−11.03*)	0.552(22.50*)	−0.180(−6.18*)
$Z2(t-1)$ (NY-MW)	0.438(17.22*)	−0.407(−15.74*)	−0.347(−11.30*)
F-Stats: ΔIBMMW	7.51*	0.14	0.09
ΔIBMPC	2.81	4.64*	5.02*
ΔIBMNY	3.32	7.83*	18.48*
$Z1$, $Z2$	160.30*	261.52*	401.69*
R^2	0.077	0.102	0.151
DW	2.000	2.000	2.000

[a]Estimates(*t*-statistics) are indicated for each variable.
*Indicates significance at the 1.0-percent level.

The procedures used in obtaining the error correction results for the MINSPAN data are similar to those in Table 3. The system lag length is $p = 2$ (see Table 2). The estimation results above are qualitatively identical to those in Table 3.

C. Ordered Tuple Results

The crucial remaining question is whether the error-correcting price adjustments are an artifact of the ordering of the data within tuples. If systematic patterns of bid-ask bounce across markets have been mistaken for error correction, the three ordered tuple data sets would yield qualitatively different error correction results than the REPLACE ALL data. Tables 5 and 6 show that this is not the case. Table 5 shows that, for each data set, the maximal eigenvalue test again indicates two cointegrating vectors, the parameters of which sum to zero. More importantly, 16 of the 18 error correction terms in the three ECM models (Table 6) have the same sign as in the original model and are significant at the 1-percent level.[11] Finally, the relative magnitudes of the coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ are qualitatively similar across the models, and the F-statistics again show the dominant effect of $Z1_{t-1}$ and $Z2_{t-1}$ (as opposed to the lagged price terms) in explaining variation.

TABLE 5

Cointegration Test Statistics for Three Ordered Tuple Data Sets

Panel A. Midwest Exchange First (T = 12,142)					
	Eigenvalues ($\hat{\lambda}$)				
	(0.148	0.114	0.00003)		
	Eigenvectors ($\hat{V}$)			H_0	$\hat{\lambda}_{max}$
IBMMW	227.683	−7.250	−0.552	$r \leq 2$	0.3
IBMPC	−193.064	−114.233	−0.280	$r \leq 1$	1,464.9
IBMNY	−34.621	121.484	−0.168	$r = 0$	1,940.7
Panel B. Pacific Exchange First (T = 11,380)					
	Eigenvalues ($\hat{\lambda}$)				
	(0.185	0.155	0.00002)		
	Eigenvectors ($\hat{V}$)			H_0	$\hat{\lambda}_{max}$
IBMMW	195.079	−102.370	−0.403	$r \leq 2$	0.3
IBMPC	−188.509	−56.759	−0.307	$r \leq 1$	1,922.5
IBMNY	−6.571	159.130	−0.289	$r = 0$	2,323.6
Panel C. New York Exchange First (T = 7,024)					
	Eigenvalues ($\hat{\lambda}$)				
	(0.153	0.106	0.00008)		
	Eigenvectors ($\hat{\lambda}$)			H_0	$\hat{\lambda}_{max}$
IBMMW	225.502	−67.948	−1.183	$r \leq 2$	0.6
IBMPC	−211.908	−72.539	0.156	$r \leq 1$	788.5
IBMNY	−13.597	140.490	0.026	$r = 0$	1,162.3

In Tables 3 and 4, a systematic pattern of bid–ask bounce across the three markets could be mistaken for error correction. To address this potential bias, the data were recollected into three subsets of the REPLACE ALL data: tuples with New York first, Midwest first, and Pacific first. The mean SPANs for these data sets are reported in Table 1. The cointegration results above show that, for each data set, the maximal eigenvalue test again indicates two cointegrating vectors. Also as before, the parameters of the cointegrating vectors sum to approximately zero.

[11] As for the other two error correction terms, one is the same sign and significant at the 10-percent level ($Z1_{t-1}$ in ΔIBMNY, Pacific First), whereas the other is insignificant at all levels ($Z1_{t-1}$ in ΔIBMNY, New York First). To save space, Table 6 suppresses the parameter estimates on the lagged price terms.

TABLE 6

Summary Estimation Results for Three Ordered Tuple Data Sets

	Dependent Variable		
	ΔIBMMW	ΔIBMPC	ΔIBMNY
Panel A. Midwest Exchange First			
Estimates(t-statistics):			
$Z1(t-1)$	−0.459(−13.09*)	0.397(10.53*)	−0.294(−6.68*)
$Z2(t-1)$	0.627(17.50*)	−0.256(−6.64*)	−0.131(−2.92*)
CONSTANT	−0.19E-04(−1.12)	−0.19E-04(−1.04)	−0.21E-04(−0.95)
F-Stats:			
ΔIBMMW	0.59	0.93	1.50
ΔIBMPC	0.38	8.92*	1.68
ΔIBMNY	2.09	4.71*	13.65*
$Z1$, $Z2$	158.95*	65.23*	156.25*
R^2	0.137	0.093	0.151
DW	2.001	2.000	2.004
Panel B. Pacific Exchange First			
Estimates(t-statistics):			
$Z1(t-1)$	−0.090(−2.58*)	0.782(23.37*)	−0.063(1.65)
$Z2(t-1)$	0.313(8.06*)	−0.581(−15.48*)	−0.422(−9.88*)
CONSTANT	−0.20E-04(−1.06)	−0.20E-04(−1.10)	−0.21E-04(−1.01)
F-Stats:			
ΔIBMMW	3.42*	1.57	1.32
ΔIBMPC	1.13	4.63*	1.57
ΔIBMNY	3.15	6.80*	2.81
$Z1$, $Z2$	43.85*	275.52*	131.22*
R^2	0.053	0.197	0.101
DW	2.000	2.000	2.001
Panel C. New York Exchange First			
Estimates(t-statistics):			
$Z1(t-1)$	−0.136(−2.15)	0.712(11.08*)	0.0004(0.006)
$Z2(t-1)$	0.240(3.61*)	−0.632(−9.35*)	−0.44(−6.19*)
CONSTANT	−0.32E-04(−1.07)	−0.33E-04(−1.06)	−0.32E-04(−1.02)
F-Stats:			
ΔIBMMW	3.69*	3.24*	0.75
ΔIBMPC	1.63	3.97*	2.70
ΔIBMNY	3.59*	4.24*	2.62
$Z1$, $Z2$	7.58*	61.46*	59.68*
R^2	0.038	0.081	0.069
DW	2.000	2.000	2.000

*Indicates significance at the 1.0-percent level.

The procedures used in estimating the above three models are similar to those in Tables 3 and 4. Using these three alternative data sets, the estimation results show that 16 of the 18 error correction terms have the same sign as in the original model in Table 3 and are significant at the 1-percent level. The relative magnitudes of the estimated coefficients on $Z1_{t-1}$ and $Z2_{t-1}$ are also qualitatively similar. Finally, the F-statistics again show the dominant effect of the $Z1_{t-1}$ and $Z2_{t-1}$ terms in explaining variation, compared to the lagged price terms.

V. Conclusion

Previous research by Lee (1993) and Blume and Goldstein (1991) implies that price discovery occurs primarily on the NYSE. This paper demonstrates that equilibrium IBM prices are also established by information revealed on the Midwest and Pacific Exchanges. All three markets react to the independent information reflected in each exchange's prices. Specifically, price changes on all three exchanges depend on price differentials relative to the NYSE. The NYSE's adjustment to a price differential with either regional exchange is smaller, on average, than the regional exchange adjustment. However, NYSE prices react to a greater degree than either Midwest or Pacific prices, when both regional prices are above or below New York.

Additional research has begun to address the relative informativeness of trades on the various exchanges and the competition for order flow between them. Using a modification of Hasbrouck's (1991), (1993) method for measuring the innovation variance in prices and volumes, Gardner and Subrahmanyam (1994) find that less informed trades are executed on the regional exchanges than on the NYSE. Our recent paper agrees with that conclusion and traces the effect of a larger asymmetric information component of the NYSE spread on the dispersion of order flow across the NYSE, the regionals, and NASDAQ (Harris, McInish, and Wood, 1994). Controlling for many other factors, in stocks and on days when trades on the NYSE are more asymmetrically informed than those on another exchange, order flow (especially by institutional traders) appears to be dispersed.[12] This may explain why the present paper detects in the error correction dynamics a statistically significant role for the satellite exchanges in the price discovery process.

[12]Our speculation concerning the role of institutional trading strategy is based on anecdotal evidence. Keim and Madhavan (1995) have begun to analyze related questions with a rich data set on institutional equity trades.

References

Back, K. "Insider Trading in Continuous Time." *Review of Financial Studies,* 5 (3, 1992), 387–409.

_______________. "Asymmetric Information and Options." *Review of Financial Studies,* 6 (3, 1993), 435–472.

Blume, M. E., and M. Goldstein. "Execution Costs on the NYSE, the Regionals and the NASD." Working Paper, Univ. of Pennsylvania (1991).

_______________. "Displayed and Effective Spreads by Market." Working Paper, Univ. of Pennsylvania (1993).

Brenner, R., and K. Kroner. "Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets." *Journal of Financial and Quantitative Analysis,* 30 (March 1995), 23–42.

Bronfman, C., and R. A. Schwartz. "Order Placement and Price Discovery in a Securities Market." Working Paper, Stern School, New York Univ. (1990).

Dickey, D., and W. A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association,* 74 (June 1979), 423–431.

_______________. "Likelihood Ratio Statistics for Auto-Regressive Time Series with a Unit Root." *Econometrica,* 49 (June 1981), 1057–1072.

Domowitz, I. "Price Discovery in Automated Trading Systems: A Global Comparison." Working Paper, Northwestern Univ. (1992).

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

Garbade, K. D., and W. L. Silber. "Dominant and Satellite Markets: A Study of Dually-Traded Securities." *Review of Economics and Statistics,* 61 (Aug. 1979), 455–460.

Gardner, A., and A. Subrahmanyam. "Multi-Market Trading and the Informativeness of Stock Trades: An Empirical Intraday Analysis." Working Paper, Columbia Univ. (1994).

Glosten, L., and P. Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics,* 14 (March 1985), 71–100.

Granger, C. W. J. "Developments in the Study of Cointegrated Economic Variables." *Oxford Bulletin of Economics and Statistics,* 48 (Aug. 1986), 213–227.

Handa, P., and R. A. Schwartz. "Dynamics of Price Discovery." Working Paper, Stern School, New York Univ. (1991).

Harris, F. H. deB.; T. H. McInish; and R. A. Wood. "The Competition for Order Flow in NYSE-Listed Stocks." Working Paper, Babcock School, Wake Forest Univ. (1994).

Hasbrouck, J. "Measuring the Information Content of Stock Trades." *Journal of Finance,* 46 (March 1991), 179–208.

_______________. "One Security, Many Markets: Determining the Contributions to Price Discovery." Working Paper, Stern School, New York Univ. (1993).

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

Johansen, S., and K. Juselius. "Maximum Likelihood Estimation and Inference on Cointegration–With Applications to the Demand for Money." *Oxford Bulletin of Economics and Statistics,* 52 (May 1990), 169–210.

Keim, D., and A. Madhavan. "Execution Costs and Investment Performance: An Empirical Analysis of Institutional Equity Trades." Working Paper, Wharton School, Univ. of Pennsylvania (1995).

Kyle, A. S. "Informed Speculation with Imperfect Competition." *Review of Economic Studies,* 56 (July 1989), 317—356.

Lee, C. "Market Integration and Price Execution for NYSE-Listed Securities." *Journal of Finance,* 48 (July 1993), 1009–1038.

McInish, T. H., and R. A. Wood. "An Analysis of Intraday Patterns in Bid-Ask Spreads." *Journal of Finance,* 47 (June 1992a), 753–764.

_______________. "Price Discovery, Volume and Regional/Third Market Trading." Working Paper, Univ. of Memphis (1992b).

Schreiber, P. S., and R. A. Schwartz. "Price Discovery in Securities Markets." *Journal of Portfolio Management,* 12 (Summer 1986), 43–48.