The Dynamics of Stock Index and Stock Index Futures Returns

Hans R. Stoll and Robert E. Whaley*

Abstract

In rational, efficiently functioning markets, the returns on stock index and stock index futures contracts should be perfectly, contemporaneously correlated. This study investigates the time series properties of 5-minute, intraday returns of stock index and stock index futures contracts, and finds that S&P 500 and MM index futures returns tend to lead stock market returns by about five minutes, on average, but occasionally as long as 10 minutes or more, even after stock index returns have been purged of infrequent trading effects; however, the effect is not completely unidirectional, with lagged stock index returns having a mild positive predictive impact on futures returns.

I. Introduction

In spite of their relatively short history, the stock index futures contract market has become a controversial topic of discussion and debate. Beginning before, but particularly since the stock market crash of October 19, 1987, stock index futures, index arbitrage, and program trading have been blamed for excessive stock market price swings. Many governmental and academic studies have examined intraday patterns of index futures and stock price changes in the days surrounding and including October 19;[1] however, little empirical analysis of the intraday comovement of the prices of index futures and stocks in more normal periods has appeared.[2]

* Owen Graduate School of Management, Vanderbilt University, Nashville, TN 37203, and The Fuqua School of Business, Duke University, Durham, NC 27706, respectively. This research was supported by a joint grant from the Chicago Board Options Exchange, the Chicago Board of Trade, and the Chicago Mercantile Exchange. Helpful comments and suggestions by seminar participants and colleagues at Duke University, The Ohio State University, Washington University, and York University are gratefully acknowledged. In particular, comments by Kal Cohen, Doug Foster, Ken French, Larry Harris, Campbell Harvey, Michael Hemler, Naoki Kishimoto, John Long, *JFQA* Managing Editor Paul Malatesta, Tom Smith, Charles Whiteman, and an anonymous *JFQA* referee greatly improved the paper. Very capable computer work was provided by Sunil Paremeswaran and John Sprow. Whaley also acknowledges the support of the Unisys Fund at The Fuqua School of Business.

[1] Detailed plots of day-by-day prices are contained in the Brady Commission Report (1988), the CFTC Report (1988), and the SEC Report (1988). A systematic examination of the relation between the price movements of the S&P 500 and its futures contract on October 19 is contained in Harris (1988).

[2] Stoll and Whaley (1986), (1987) examine minute-by-minute price change behavior of the S&P 500 and Major Market indexes in the days surrounding the expirations of the S&P 500 and MMI

The purpose of this paper is to model empirically the temporal relation between the price movements of index futures contracts and stocks. In the process, the paper provides insights on the volatility of index futures versus stock indexes and the extent to which futures overshoot true values. The paper is distinguished from prior work in several ways. First, a longer time interval—five years—and a finer return grid—five minutes—is examined than in other papers. Second, the delay in the price change of a stock index due to the infrequent trading of the component stocks is treated explicitly. A theoretical model of observed portfolio returns that incorporates the effects of infrequent trading and stock bid/ask spreads is developed and estimated, with the residuals of this model (i.e., return innovations) then used to proxy for "true" stock index returns. To verify the results, the returns of the most actively traded individual common stock—IBM—also are used as a proxy for true index returns. Third, both the Chicago Mercantile Exchange's S&P 500 and the Chicago Board of Trade's Major Market Index futures contracts are considered.

The paper is organized as follows. The theory underlying the covariation of stock index futures and stock index returns is outlined in Section II, and the effects of infrequent trading and transaction costs are identified. In Section III, the data are described. All tests are based on 5-minute, intraday returns. The investigation period begins with the introduction of the S&P 500 futures contract by the Chicago Mercantile Exchange in April 1982, extends through the introduction of the Major Market Index futures contract by the Chicago Board of Trade on July 23, 1984, and ends on March 31, 1987. Infrequent trading and bid/ask price effects are examined in Section IV. Serial correlations of returns in stock indexes, stock index futures, and IBM reveal the expected infrequent trading and bid/ask patterns. The effects of infrequent trading and bid/ask spreads on the observed structure of returns are modeled and estimated, and then return innovations are generated. In Section V, the temporal relation between the futures and stock returns is estimated in a multiple regression framework. Generally speaking, the returns in the futures market lead those in the stock market, even after adjusting for the infrequent trading of stocks. The evidence also shows that the lead has diminished through time. The paper concludes in Section VI with a summary.

II. Theory

The theoretical relation between the price of an index futures contract and the price level of the underlying index is,

$$F_t = S_t e^{(r-d)(T-t)}, \tag{1}$$

futures contracts, but do not examine nonexpiration days. Kawaller, Koch, and Koch (1987a), (1987b) use intraday data to examine price changes of the S&P 500 index and index futures; however, they examine only three quarters of the data and do not correct for infrequent trading. MacKinlay and Ramaswamy (1988) also use intraday data, but they focus on deviations from the cost of carry equilibrium. Ng (1987) uses interday data to investigate the price behavior of S&P 500 index futures prices and its ability to predict the S&P 500 index level. Chan, Chan, and Karolyi (1990) use intraday price data to examine the transmission of volatility between the stock and index futures markets.

where F_t is the index futures price at time t, S_t is the index price at time t, $r-d$ is the net cost of carrying the underlying stocks in the index, that is, the rate of interest cost r less the rate at which dividend yield accrues to the stock index portfolio holder d. T is the expiration date of the futures contract, so $T-t$ is the time remaining in the futures contract life. Note that in this formulation the riskless rate of interest and the dividend yield on the underlying stock index are assumed to be known, constant, continuous rates.

The market force driving the cost-of-carry relation (1) is the never-ending search for a "free lunch." When the futures price is above the level implied by the right-hand side of (1), a riskless arbitrage profit equal to the difference between the futures price and the index price plus the cost of carry, a long arbitrage profit of $F_t - S_t e^{(r-d)(T-t)}$ can be earned by selling the futures contract and buying the stock index portfolio, financing the stock purchase with riskless borrowings. On the other hand, when the futures price falls below the right-hand side of (1), a short arbitrage profit of $S_t e^{(r-d)(T-t)} - F_t$ can be earned by buying the futures and selling the portfolio of stocks, investing the proceeds of the sale of stock at the riskless rate of interest. The use of a single, computer-generated order to buy or sell an entire portfolio of stocks is known as "program trading."

In perfectly efficient and continuous futures and stock markets absent transaction costs, riskless arbitrage profit opportunities should not appear so the cost-of-carry relation (1) should be satisfied at every instant t during the futures contract life. If such is the case, the instantaneous rate of price appreciation in the stock index equals the net cost of carry of the stock portfolio plus the instantaneous relative price change of the futures contract; that is,

$$R_{S,t} = (r-d) + R_{F,t}, \tag{2}$$

where $R_{S,t} = \text{Ln}(S_t/S_{t-1})$ and $R_{F,t} = \text{Ln}(F_t/F_{t-1})$.

Several implications follow from (2) under the assumptions that the short-term interest rate and the dividend yield rate of the stock index are constant and that the index futures and stock markets are efficient and continuous:

(a) The expected rate of price appreciation on the stock index portfolio $E(R_{S,t})$ equals the net cost of carry $(r-d)$ plus the expected rate of return on the futures contract $E(R_{F,t})$.

(b) The standard deviation of the rate of return on the futures contract equals the standard deviation of the rate of return of the underlying stock index.

(c) The contemporaneous rates of return of the futures contract and the underlying stock portfolio are perfectly positively correlated.

(d) The rates of return of the futures contract and of the underlying stock index portfolio are serially uncorrelated.[3]

(e) The noncontemporaneous rates of return of the futures contract and the underlying stock portfolio are uncorrelated.

[3] Technically speaking, more than an assumption of market efficiency is needed to ensure serially uncorrelated rates of return. It must also be the case that the expected rates of return of the futures and stock index are constant (see Fama (1976), pp. 149–151). An assumption of constant expected returns is reasonable here since *intraday* rate of return series are examined.

Naturally, all of the above implications are based on the assumption that the cost-of-carry relation (1) holds at all points in time. It has been shown, however, that (1) does not hold exactly; indeed one of the puzzles in stock index futures is the frequency with which deviations from (1) are observed. Stoll and Whaley ((1986), Table 23A), for example, report frequent violations of the cost-of-carry relation in excess of transaction costs using hourly S&P 500 index and index futures data during the period April 1982 through December 1985. The frequency of violation is nearly 80 percent for the June 1982 futures contract; however, for more recent contract maturities, the frequency falls below 15 percent. MacKinlay and Ramaswamy ((1988), Table 6) report similar results for the S&P 500 futures contracts expiring in September 1983 through June 1987. Using 15-minute price data, they find that the cost-of-carry relation is violated 14.4 percent of the time, on average.

Violations of the cost-of-carry relation may appear for a variety of reasons. Some are purely technical. An important one is the infrequent trading of stocks within the index. Markets for individual stocks are not perfectly continuous. Consequently, stock index prices, which are averages of the last transaction prices of component stocks, lag actual developments in the stock market. Fisher (1966) describes this phenomenon. Cohen, et al. ((1986), Ch. 6) give a more general discussion of serial correlation of stock index returns in terms of delays in the price adjustment of securities. Lo and MacKinlay (1988) model the effects of infrequent trading on index returns under restrictive assumptions. Assuming that the index futures prices instantaneously reflect new information, observed futures returns should be expected to lead observed stock index returns because of infrequent trading, even though there is no economic significance to this behavior whatsoever.

A second reason for violation of relation (2) is that transaction costs tend to induce noise in the relation (2). The prices used in the computation of returns are transaction prices, and these transaction prices tend to fluctuate randomly between bid and ask levels. This random price movement between bid and ask prices in successive transactions induces negative serial correlation in observed returns even though the true returns are serially independent.[4] At the individual security level, the negative serial correlation due to the bid/ask price effect is understandable, but the effect seems less likely when one considers a stock index portfolio for which movements between the bid and ask for some stocks could be offset by opposite movements from the ask to bid for other stocks; however, to the extent that the rates of return of the stocks in the index are positively correlated and/or that the index is narrowly based, negative serial correlation in individual stock returns attributable to the bid/ask price effect also might appear in the stock index returns.

A third reason for violation of the cost-of-carry relation has to do with time delays in the computation and reporting of the stock index value. Once a transaction in the stock market takes place, the transaction information is entered into a computer and transmitted to the particular service that updates and transmits the

[4] Roll (1984) develops a simple model to show the relation between the bid/ask spread and the serial covariance in stock returns. Stoll (1989) and others have extended Roll's work.

index level.[5] Three time delays are therefore possible: (a) the delay in entering the stock transaction into the computer; (b) the delay in computing and transmitting the new index value; and, (c) the delay in recording the stock index value at the futures exchange. Assuming that new information arrives in the stock and futures markets simultaneously and that price changes in the futures market are recorded instantaneously, such delays would tend to show the futures market returns leading stock index returns.

Finally, lead/lag behavior of stock index and stock index futures returns may reflect the greater speed with which investors' views are reflected in futures markets. Investors with strong beliefs about the direction of the market as a whole (as opposed to a trend in the price of an individual stock) may trade index futures rather than individual stocks because transaction costs are lower and the degree of leverage attainable is higher. Such trading moves futures prices first, and then pulls stock prices when index arbitrage responds to the deviations from the cost-of-carry relation (1).

III. Data

A. Data Description and Sources

The data used in this study were obtained from three separate sources—the Chicago Mercantile Exchange (CME), the Chicago Board of Trade (CBOT), and Francis Emory Fitch, Inc. (''Fitch''). The CME provided the S&P 500 index and index futures price data for the period April 21, 1982, through March 31, 1987. These data, referred to as ''Quote Capture'' information, contain the time (to the nearest 10 seconds) and price (to the nearest 0.05 index points) of every futures transaction in which the price has changed from the previously recorded transaction, as well as the S&P 500 index level (to the nearest 0.01 index points) each time it is computed and transmitted to Chicago. Prior to June 13, 1986, the stock index was computed approximately once a minute; but, since that time, it has been computed and reported approximately four times per minute.[6] The S&P 500 futures are on a quarterly expiration cycle (i.e., March, June, September, and December), and it is always the nearby contract that is the most active in terms of trading volume. Since the time series tests require the most frequent return observations possible, only the data for the nearby futures contract are used.

The CBOT provided the Major Market Index (MMI) and MMI futures price data for the period July 23, 1984, through March 31, 1987. Like the Quote Capture information from the CME, the data file contains the time (to the nearest second) and price (to the nearest 0.10 index points before August 16, 1985, and

[5] Two stock indexes are considered here—the S&P 500 and the MMI. ADP Brokerage Information Services Group handles the computation and dissemination of the price level of the S&P 500 and the American Stock Exchange handles the MMI.

[6] ADP Brokerage Information Services Group computes the S&P 500 index level every time one of the index's component stocks has a price change. The number of shares used in the computation of the index are provided to ADP by Standard & Poors' once a week (each Wednesday) and are held constant through the week. As a matter of routine, ADP currently transmits the index level to the CME at a rate of four times per minute. The authors are grateful to Leo McBlain at ADP for providing this information.

to the nearest 0.05 index points from that date forward) of futures transactions[7] as well as the stock index (to the nearest 0.01 index points) information. The MMI quotes are recorded at approximately 15-second intervals.[8] The MMI futures contracts have monthly expiration dates, and only the nearby futures is used for testing.

Transaction-by-transaction data for IBM during all trading days in the years 1982 through 1986 were supplied by Fitch. The transaction records are time stamped to the nearest minute and contain both price and volume information. IBM was chosen because it is the most frequently traded stock within both the S&P 500 and Major Market indexes.

The fact that the three sources of data span three different, but overlapping, time periods causes three different samples to be used in the subsequent tests. The samples are labeled by the length of the sample periods in days:

1,249-Day Period: April 21, 1982, through March 31, 1987. The entire transaction price history of the S&P 500 Index futures contract along with its corresponding stock index value.

678-Day Period: July 23, 1984, through March 31, 1987. The entire transaction price history of the Major Market Index futures contract along with its corresponding stock index value.

609-Day Period: July 23, 1984, through December 31, 1986. The intersection of the transaction price histories of the S&P 500 Index, Major Market Index, and IBM samples.

The complete S&P 500 (1,249-day) and MMI (678-day) price histories are used where comparisons of the futures and its underlying index are made. The 609-day period common to the S&P 500, MMI, and IBM samples is used where comparisons of the index and index futures with IBM are informative.

Finally, in the section devoted to assessing index futures market maturation effects, a measure of daily stock market activity is necesssary. The proxy chosen is the total number of trades for all stocks on the NYSE each day. These data were obtained from the NYSE.

B. Return Series

None of the transaction price series, including those of the stock indexes, has price observations uniformly spaced in time.[9] It is, therefore, necessary to convert the transaction-by-transaction prices to returns over a fixed time interval. While an interval as short as one minute is feasible, a 5-minute interval is chosen

[7] The original MMI futures contract was denominated as 100 times the index value and had minimum price increments of 0.10 index points. On July 7, 1985, the CBOT introduced a second futures contract, the Maxi MMI, which is denominated as 250 times the index value and has minimum price increments of 0.05 index points. From July 7, 1985, through September 19, 1986, both futures contracts traded simultaneously; however, the larger contract quickly dominated in terms of trading volume. On September 19, 1986, the smaller contract was discontinued. In terms of splitting the MMI sample so as always to include only the most active MMI futures contract, August 16, 1985, the expiration date of the August 1985 futures contract, was used.

[8] AMEX computes and disseminates the MMI level at 15-second intervals. The time stamps that appear on the CBOT data base are the times at which the prices are received and recorded by the CBOT. The authors are grateful to Charles Faurot at the AMEX for explaining the transmission process.

[9] The word "transaction" is used to describe the stock index quote for ease in exposition.

to mitigate the effects of the errors-in-the-variables problem induced by nonsimultaneous futures and index price observations. Each trading day is partitioned into 5-minute intervals, beginning with the opening of the NYSE—at 9:00 AM (CST) before September 30, 1985, and at 8:30 AM (CST) thereafter. The first index futures price and the first stock index value observed in the 5-minute interval are recorded. The first 5-minute interval of the day is skipped if it does not contain price observations on both series, and the next 5-minute interval is searched until an interval with both prices is found. Thereafter, every 5-minute interval is used even if both prices are not available. The S&P 500 index (futures), for example, does not have prices reported in 0.04 (0.27) percent of the 5-minute intervals.

During the 1,249-day sample period, the median time elapsed between the grid time and the first transaction after the grid time is 26 seconds for the S&P 500 index and 10 seconds for the S&P 500 futures contract. Although these prices are classified as being contemporaneous in the empirical investigations to follow, the index observation generally occurs after the futures transaction, so, holding other factors constant, a bias in favor of finding index returns leading futures returns is introduced. For the 678-day sample period, the median times for the MM index and the MMI futures contract are 7 and 27 seconds, respectively. Here the bias is in favor of finding the futures leading the index. Finally, the median time elapsed for IBM is 0 seconds. This value is not very precise because IBM transactions are reported only to the nearest minute. What this figure implies is that the median time is less than 30 seconds. Without a more precise estimate, it is impossible to say how the temporal relation test results in the regressions using IBM returns are affected.

The 5-minute price series are then used to generate the time series of instantaneous rates of return. The returns for the futures contract and the stock index are defined as $R_{F,t} = 100*\text{Ln}(F_t/F_{t-1})$ and $R_{S,t} = 100*\text{Ln}(S_t/S_{t-1})$, respectively. The last interval of the day ends at 3:00 PM (CST), so a maximum of 73 5-minute price observations (72 returns) are possible for each day prior to September 30, 1985, and a maximum of 79 (78 returns) for each day after that date.[10] In all, five rate of return series are created using the same procedure. The return series are for: (a) S&P 500 Index, (b) the S&P 500 index futures, (c) the MM Index, (d) the MMI futures, and (e) IBM. On each day, each return series begins with the 5-minute interval that contains a price for that series and for the S&P 500 index futures series. Because returns are computed within each day using only intraday prices, overnight returns are not included in any of the series.

Also excluded from the analysis are the first two 5-minute rates of return of each stock index series each day. The reason for this exclusion is that at the beginning of the day the index values are computed using, for the most part, closing stock prices from the previous day. To the extent that closing prices are inaccurate (or stale) reflections of opening values, there may be noise in the index level and, hence, in index returns, until all stocks within the index have traded. Stoll and Whaley (1990) report that the average time to open for stocks in

[10] The S&P 500 and MM index futures markets close at 3:15 PM (CST), while the NYSE closes at 3:00 PM. Since the focus of the study is on the temporal relation between returns in the two markets where simultaneous price observations are necessary, futures returns after 3:00 PM are ignored.

the S&P 500 index (average time elapsed between exchange opening and opening transaction) is between five and seven minutes. Disregarding the first two 5-minute return observations each day, therefore, mitigates the effects of stale price information.

IV. Infrequent Trading and Bid/Ask Price Effects

The objective of this paper is to estimate the empirical relation between index futures returns and the returns of the underlying index. Before doing so, however, it is necessary to consider the univariate time series properties of the observed stock index return series. Observed index portfolio returns are not accurate reflections of "true" index returns because not all stocks in the index trade in every interval of time and because index levels are based on transaction prices of individual stocks, not true prices. To proxy for the true but unobserved index returns, two approaches are used. First, the effects of infrequent trading and the bid/ask spread on the observed structure of index returns are modeled and estimated. Deviations from the estimated model (i.e., the return innovations) are then used as instruments for the true index returns. Second, the returns of a single stock, IBM, are used as a proxy for the true index returns. IBM was chosen based on its presence in both the S&P 500 and MM indexes, its large market capitalization, and its highly active secondary market. In particular, its highly active secondary market should mitigate, if not entirely eliminate, the effects of infrequent trading.

This section deals with the effects of infrequent trading and the bid/ask spread on observed returns. First, serial correlations of the observed return series are examined. The expected positive serial dependence in the returns of the stock indexes and the expected negative serial dependence in individual stock returns appear. Second, a theoretical model describing the effects of infrequent trading and the bid/ask spread on observed stock portfolio returns is developed. Finally, the theoretical model's parameters are estimated for the S&P 500, MMI, and IBM return series and the return innovations are generated. The estimated model appears to control for the effects of infrequent trading and bid/ask spreads very well.

A. Serial Correlation in Observed Returns

Serial correlations are estimated for lags one through 12, that is, up to one hour of trading time, using observed 5-minute returns across the 1,249-day and 609-day sample periods. As noted earlier, overnight returns are excluded, so 1,249 or 609 return observations, depending on the sample period, are lost each time the order of the serial correlation k increases. Thus, serial correlation estimates are never contaminated by using returns from adjacent days. For the stock indexes, the first two returns each day also are excluded. The serial correlation estimates are reported in Table 1.

In Table 1, note first that the lag one serial correlation coefficient for IBM is reasonably large and significantly negative, -0.0579. This is the pattern predicted by models of the bid/ask price effect. The effect should not persist at higher order lags unless IBM shares do not trade in every 5-minute interval of

TABLE 1

Estimated Serial Correlation Coefficients of Observed Returns of the S&P 500 Index (R_S^O), the S&P 500 Index Futures Contract (R_F^O), the Major Market Index (R_M^O), and IBM (R_I^O)

	S&P 500 Index (*S* and *F*)					MM Index (*M*)			IBM (*I*)		
		$\rho_k(R^O_{S,t}, R^O_{S,t-k})$		$\rho_k(R^O_{F,t}, R^O_{F,t-k})$			$\rho_k(R^O_{M,t}, R^O_{M,t-k})$			$\rho_k(R^O_{I,t}, R^O_{I,t-k})$	
Lag k	No. of Obs.[a]	$\hat{\rho}_k$[b]	$t(\hat{\rho}_k)$[c]	$\hat{\rho}_k$[b]	$t(\hat{\rho}_k)$[c]	No. of Obs.[a]	$\hat{\rho}_k$[b]	$t(\hat{\rho}_k)$[c]	No. of Obs.[a]	$\hat{\rho}_k$[b]	$t(\hat{\rho}_k)$[c]
609-Day Period: July 23, 1984–Dec. 31, 1986											
1	43,193	0.4477	104.07	0.0069	1.44	43,083	0.2443	52.29	43,978	−0.0579	−12.16
2	42,584	0.1428	29.78	−0.0165	−3.40	42,474	0.0185	3.82	43,369	−0.0423	−8.81
3	41,975	0.0221	4.54	−0.0256	−5.24	41,865	−0.0473	−9.70	42,760	−0.0259	−5.35
4	41,366	0.0128	2.60	−0.0039	−0.79	41,256	−0.0212	−4.31	42,151	0.0003	0.07
5	40,757	0.0233	4.70	−0.0013	−0.27	40,647	−0.0050	−1.01	41,542	0.0136	2.77
6	40,148	0.0146	2.93	0.0055	1.10	40,038	−0.0105	−2.09	40,933	0.0103	2.09
7	39,539	0.0131	2.61	0.0049	0.97	39,429	−0.0113	−2.24	40,324	0.0131	2.62
8	38,930	0.0250	4.94	0.0131	2.59	38,820	0.0038	0.75	39,715	0.0113	2.25
9	38,321	0.0233	4.57	0.0160	3.13	38,211	0.0101	1.98	39,106	0.0128	2.52
10	37,712	0.0123	2.39	0.0039	0.75	37,602	0.0053	1.03	38,497	0.0170	3.34
11	37,103	0.0080	1.54	0.0076	1.47	36,993	0.0060	1.15	37,888	0.0031	0.61
12	36,494	0.0109	2.07	0.0007	0.13	36,384	0.0091	1.74	37,279	0.0104	2.00
1249-Day Period: April 21, 1982–March 31, 1987											
1	86,952	0.5117	175.61	0.0229	6.77						
2	85,703	0.2654	80.60	0.0265	7.76						
3	84,454	0.1312	38.46	0.0015	0.45						
4	83,205	0.0759	21.96	−0.0137	−3.96						
5	81,956	0.0460	13.17	−0.0222	−6.36						
6	80,707	0.0199	5.64	−0.0108	−3.06						
7	79,458	0.0077	2.18	−0.0087	−2.46						
8	78,209	0.0154	4.32	−0.0015	−0.42						
9	76,960	0.0195	5.42	0.0039	1.07						
10	75,711	0.0110	3.04	−0.0030	−0.83						
11	74,462	0.0018	0.49	0.0047	1.29						
12	73,213	0.0019	0.51	0.0002	0.07						

[a] The number of observations used in the computation of the serial correlation coefficient. Note that as the lag k is incremented by one, the number of observations lost equals the number of days in the sample period. This reflects the loss of one return each day of the sample. The serial correlation coefficient estimates, therefore, are not contaminated by using returns from adjacent days.

[b] The estimated lag k serial correlation coefficient across all 5-minute returns in all days of the period excluding overnight returns. For the stock indexes, the first two returns each day are excluded.

[c] The t-ratio corresponding to the null hypothesis ρ_k equals zero.

time. The lingering negative, but less significant, serial correlation at lags 2 and 3, −0.0423 and −0.0259, suggests that occasionally there are periods during the trading day when IBM does not trade for 10 minutes or more; however, the relative magnitudes of the coefficient estimates are somewhat misleading. An examination of IBM's transactions during the 609-day sample period revealed that IBM's stock traded in more than 98 percent of the 44,587 5-minute intervals in the sample. And, of the intervals in which IBM did not, there were only 125 instances in which IBM did not trade in two or more consecutive 5-minute intervals.

As predicted by models of infrequent trading, the stock index return series exhibit strong positive serial correlation. The infrequent trading effect is stronger for the broad-based S&P 500 index than the Major Market Index, which consists of 20 highly active stocks. The lag one coefficient is 0.4477 for the S&P 500 index, compared with 0.2443 for the MMI during the 609-day sample period.

Moreover, the positive serial dependence disappears after lag one for the MMI, while it persists through lag two for the S&P 500. Note also that the lag three serial correlation coefficient of the returns of the MMI is negative and reasonably large, −0.0473. This result may be due to the bid/ask price effect. If there are periods during the trading day in which the stocks in the MMI do not trade, the negative serial correlation in the returns of the stocks within the index may appear in the index returns because the MMI is narrowly based. If the index has a large base such as the S&P 500, the bid/ask price effects in the individual stock returns tend to disappear in the index portfolio returns as a result of diversification. The MMI contains only 20 stocks, however, and the apparent negative serial dependence is probably due to individual stock bid/ask spreads.

The serial correlation coefficients for the returns of the S&P 500 futures contracts, on the other hand, are negligible at all lags. Since the futures returns are for a single financial instrument rather than a portfolio of securities, no positive serial dependence due to infrequent trading should appear. Negative serial dependence resulting from the bid/ask effect is possible; however, the coefficient estimate at lag one during the 609-day sample period is positive and insignificant, 0.0069. The lag one coefficient in the overall 1,249-day sample, 0.0229, is also very small, although its t-ratio is 6.77. Because of the large number of observations, very small serial correlations can be statistically significant under the null hypothesis of zero serial correlation, even though other specific null hypotheses also may be rejected. In the interpretation of the serial correlation results as well as the regression results to follow, significance is evaluated in economic terms, and the magnitudes of the first-order serial correlation in futures returns indicate clearly that the bid/ask spread in the futures market contributes little to the variability of the S&P 500 index futures returns.[11] Futures returns appear to have little or no memory.

Finally, it is worthwhile to note that the number of lags in which meaningful positive serial correlation appears for the S&P 500 index returns is greater for the overall 1,249-day sample period than for the more recent 609-day sample period. Where lags one through four are important in the second panel, only lags one and two are important in the first panel. This evidence indirectly supports the fact that trading volume in the stock market has increased dramatically in recent years. Nonetheless, because returns are being measured over such a short interval of time, infrequent trading remains a problem and needs to be modeled before investigating the temporal relation of returns in the futures and stock markets. The bid/ask price effect, although smaller in magnitude, produces significant negative serial correlation in both the return series of IBM and the MM index and also needs to be considered explicitly in modeling the time series of observed returns.

B. Modeling Infrequent Trading and Bid/Ask Price Effects

The effects of infrequent trading and bid/ask price movements on stock

[11] During the investigation period, the typical bid/ask spread of the nearby S&P 500 futures contract was about 0.05, while the futures price was about 200. This implies a relative bid/ask spread of less than 0.1 of 1 percent. This compares with an average relative spread of nearly 1.5 percent for NYSE stocks in 1979. See Stoll and Whaley (1983), pp. 73.

portfolio returns may be assessed in a variety of ways.[12] The model used here is simple, yet general, and is developed in two separate steps. First, we model the effects of the bid/ask spread on stock and stock portfolio return series assuming that stocks trade in every interval of time. Second, we extend the model by allowing stocks to trade less frequently, that is, at least once every n periods. Throughout this subsection, we assume that stock returns are independent and identically distributed through time. The return on stock i in period t is defined as,

$$R_{i,t} = \mu_i + \eta_{i,t}\,, \tag{3}$$

where μ_i is the expected return of stock i, and $\eta_{i,t}$ is the mean-zero return innovation of stock i in period t.

1. Bid/Ask Price Effects

In the presence of a bid/ask price effect and under the assumption that *a stock trades at least once during every interval of time,* the observed return of stock i in period t may be written,

$$R^*_{i,t} = \mu_i + \eta_{i,t} + \vartheta_{i,t} - \delta_i\vartheta_{i,t-1}\,, \tag{4}$$

where ϑ_t is a mean zero, i.i.d. disturbance. The intuition underlying the $MA(1)$ process governing ϑ_t in (4) is that, in a given period, the observed return equals the true return plus the sum of two bid-ask price errors—one at the beginning and one at the end of the period over which the return is computed. Under the assumption that *all stocks within a portfolio trade at least once during every interval of time,* the observed return of stock portfolio S in period t may be written,

$$\begin{aligned} R^*_{S,t} &= \sum_{i=1}^{m} X_i R^*_{i,t} \\ &= \sum_{i=1}^{m} X_i\left(\mu_i + \eta_{i,t} + \vartheta_{i,t} - \delta_i\vartheta_{i,t-k}\right) \\ &= \mu_S + \eta_{S,t} + \sum_{i=1}^{m} X_i\left(\vartheta_{i,t} - \delta_i\vartheta_{i,t-k}\right)\,, \end{aligned} \tag{5}$$

where X_i is the proportion of the market value of the portfolio accounted for by the i-th stock, m is the number of stocks in the portfolio, and the following notational substitutions have been made: $\mu_S \equiv \Sigma_{i=1}^{m} X_i\mu_i$, and $\eta_{S,t} \equiv \Sigma_{i=1}^{m} X_i\eta_{i,t}$. Equation (5) states that the observed portfolio return is the sum of the expected portfolio return, the portfolio return innovation in period t, and an error compo-

[12] Scholes and Williams (1977) examine the effects of nonsynchronous prices on the estimation of relative systematic risk. Using a similar framework, Dimson (1979), Lo and MacKinlay ((1988), pp. 56–60), and Muthuswamy (1989) model the effect of infrequent trading (and/or nonsynchronous price) on stock index returns. Stephan and Whaley (1990) model the effect of the bid/ask spread on stock price changes. Cohen, Maier, Schwartz, and Whitcomb ((1986), pp. 114–120) model both the infrequent trading and bid/ask price effects in a market model framework.

nent represented as a weighted average of stock-specific moving average processes. Assuming the bid/ask errors in period t ($\vartheta_{i,t}$'s) are independent across stocks and the X_i's are of order $1/m$, the error component disappears and $R^*_{S,t} \approx \mu_S + \eta_{S,t}$ as the number of stocks in the portfolio increases; however, for portfolios with relatively few stocks, the error component may be nonzero.

2. Infrequent Trading Effects

Equation (5) is a model of observed portfolio returns in which all stocks are known to trade at least once every interval of time; however, for time intervals as short as five minutes (the time interval used in the empirical analyses that follow), not all stocks trade every interval or even every second interval. To generalize the model, we assume that all stocks trade at least once every n intervals. If *all stocks within the portfolio trade at least once every n intervals of time,* the underlying observed stock portfolio return in period t ($R^o_{S,t}$) may be expressed as a weighted average of the contemporaneous and lagged portfolio returns from Equation (5), ($R^*_{S,t-k}$, $k = 0, 1, 2, \ldots, n-1$), plus a random mean-zero error disturbance $\nu_{S,t}$; that is,

$$(6) \qquad R^o_{S,t} = \sum_{k=0}^{n-1} \omega_{S,k} R^*_{S,t-k} + \nu_{S,t} \,.$$

The weights, $\omega_{S,k}$, are assumed to (a) be positive and constant through time, (b) decline with k, and (c) sum to one. The constraint $\Sigma_{k=0}^{n-1} \omega_{S,k} = 1$ ensures that each observed portfolio return is fully reflected in contemporaneous and future observed returns. The weights are subscripted by "S" to denote that they are portfolio specific, that is, the $\omega_{S,k}$, $k = 0, \ldots, n-1$'s are different for different portfolios (S). The disturbance term $\upsilon_{S,t}$ is assumed to be independent and identically distributed through time.

One interpretation of the formulation (6) is that only a fraction $\omega_{S,0}$ of the portfolio's true return in period t is observed in period t, where the fraction depends upon the proportion of stocks within the portfolio that trade every period as well as the proportion of the market value of the portfolio that those stocks constitute. If all stocks trade in every 5-minute interval, $\omega_{S,0} = 1$ and all other weights are zero. Fraction $\omega_{S,1}$ of the portfolio return in period t is not observed until period $t+1$ and is attributable to stocks that trade every two periods on average. Fraction $\omega_{S,2}$ is not observed until $t+2$, and so on through the next n periods, when the return attributable to the most inactively traded stock in the portfolio (which trades every n periods) is reflected in the observed portfolio return.

Under the above assumptions, the observed portfolio return in period t may be written as a function of lagged observed portfolio returns, that is,

$$(7) \qquad R^o_{S,t} = \omega_{S,0} R^*_{S,t} + \sum_{k=1}^{\infty} \gamma_{S,k} R^o_{S,t-k} + \nu_{S,t} - \sum_{k=1}^{\infty} \gamma_{S,k} \nu_{S,t-k} \,,$$

where the $\gamma_{S,k}$'s are functions of the $\omega_{S,k}$'s in (6). (Appendix A contains the proof.) In general, the values of the $\gamma_{S,k}$'s may be positive or negative but decline in absolute value as k increases (due to the assumption that $\omega_{S,0} > \omega_{S,1} > \ldots >$

$\omega_{S,n-1})$. Equation (7) shows that observed stock portfolio returns follow an *ARMA*(p,q) process with a random intercept, where the coefficients on the lagged autoregressive and moving average parameters are the same and the order of the process is infinite.

To complete the model, we must explicitly recognize the effects of bid/ask pricing errors, so we substitute (5) into (7) to get,

$$
\begin{aligned}
R^o_{S,t} &= \omega_{S,0}\left[\mu_S + \eta_{S,t} + \sum_{i=1}^{m} X_i\left(\vartheta_{i,t} - \delta_i\vartheta_{i,t-k}\right)\right] \\
&\qquad + \sum_{k=1}^{\infty}\gamma_{S,k}R^o_{S,t-k} + \nu_{S,t} - \sum_{k=1}^{\infty}\gamma_{S,k}\nu_{S,t-k} \qquad (8) \\
&= \omega_{S,0}\mu_S + \sum_{k=1}^{\infty}\gamma_k R^o_{S,t-k} + \omega_{S,0}\left(\eta_{S,t} + \vartheta_{S,t}\right) + \nu_{S,t} \\
&\qquad - \gamma_1\nu_{S,t-1} - \omega_{S,0}\sum_{i=1}^{m}X_i\delta_i\vartheta_{i,t-1} - \sum_{k=2}^{\infty}\gamma_k\nu_{S,t-k} \\
&= \omega_{S,0}\mu_S + \sum_{k=1}^{\infty}\gamma_k R^o_{S,t-k} + \epsilon^*_{S,t} - \gamma_1\epsilon^{**}_{S,t-1} - \sum_{k=2}^{\infty}\gamma_k\epsilon^{***}_{S,t-k} \\
&\approx \omega_{S,0}\mu_S + \sum_{k=1}^{\infty}\phi_k R^o_{S,t-k} + \epsilon_{S,t} - \sum_{k=1}^{\infty}\theta_k\epsilon_{S,t-k},
\end{aligned}
$$

where $\vartheta_{S,t} \equiv \Sigma_{i=1}^{m}X_i\vartheta_{i,t}$. In the development of (8), the disturbance terms $\nu_{S,t}$, $\vartheta_{i,t}$, and $\eta_{S,t}$ are aggregated to form the terms $\epsilon^*_{S,t}$, $\epsilon^{**}_{S,t-1}$, and $\epsilon^{***}_{S,t-k}$. Given the assumptions that $\nu_{S,t}$, $\vartheta_{i,t}$, and $\eta_{S,t}$ are mean-zero and i.i.d., $\epsilon^*_{S,t}$, $\epsilon^{**}_{S,t-1}$, and $\epsilon^{***}_{S,t-k}$ are also mean-zero and independent, but they do not necessarily share the same variance. Hence, the transition from the second to the last line to the last line of (8) invokes an assumption of homoskedasticity.

Equation (8) shows that, when the effects of both infrequent trading and the bid/ask spread are incorporated, observed portfolio returns follow an *ARMA*(p,q) process of infinite order. The error term $\epsilon_{S,t}$ contains three error components: (a) $\nu_{S,t}$, the random error from the infrequent trading model (6), (b) $\vartheta_{S,t}$, a weighted-average error from the individual stock bid/ask spreads, and (c) $\eta_{S,t}$, the true return innovation in the stock portfolio S in period t. In the absence of infrequent trading and bid/ask price effects, $\epsilon_{S,t} = \eta_{S,t}$. Hence, the error term measures the true return innovation in the stock portfolio return in period t. In the presence of infrequent trading and bid/ask price effects, the observed return innovation $\epsilon_{S,t}$ is a noisy but unbiased proxy for the true return innovation $\eta_{S,t}$ and is used in the investigations of the temporal relation between the index futures and stock market returns in the next section.

C. Estimating Infrequent Trading and Bid/Ask Price Effects

The effects of infrequent trading and the bid/ask spread have been shown to cause observed portfolio returns to follow an *ARMA*(p,q) process. In this subsection, the parameters of *ARMA* model are estimated. Since trading activity varies by day, the model is estimated each day for each return series. The order of the model is the same across days of the sample period. For IBM's returns, only the bid/ask price effect needs to be modeled.[13] The serial correlation estimates reported for IBM in Table 1 indicate that IBM trades once every five minutes or less, on average, but occasionally does not trade for 10 or 15 minutes, so an *MA*(3) model is most appropriate. For the S&P 500 and MM stock indexes, an *ARMA*(2,3) was used. Other *ARMA* specifications also were estimated; however, more parsimonious models did not perform as well in eliminating meaningful serial correlation in the residuals. The average daily parameter estimates of each of the fitted models are reported in Table 2, and the serial correlation estimates of the observed return innovations are reported in Table 3.

TABLE 2

Parameter Estimates from *ARMA*(*p*,*q*) Regressions Using Stock and Stock Index Returns

$$R^o_{i,t} = \mu + \phi_1 R^o_{i,t-1} + \phi_2 R^o_{i,t-2} + \epsilon_{i,t} - \theta_1\epsilon_{i,t-1} - \theta_2\epsilon_{i,t-2} - \theta_3\epsilon_{i,t-3} \quad i = S, M$$
$$R^o_{I,t} = \mu + \epsilon_{I,t} - \theta_1\epsilon_{I,t-1} - \theta_2\epsilon_{I,t-2} - \theta_3\epsilon_{I,t-3}$$

	S&P 500 Index (*S*)		MM Index (*M*)		IBM (*I*)	
Parameter	Average Parameter Estimate[a]	Standard Error[b]	Average Parameter Estimate[a]	Standard Error[b]	Average Parameter Estimate[a]	Standard Error[b]
609-Day Period: July 23, 1984–Dec. 31, 1986						
$\bar{R}^2$	0.1747		0.0935		0.0265	
$\hat{\mu}$	0.0003	0.0003	0.0006	0.0005	0.0000	0.0006
$\hat{\phi}_1$	0.4711	0.0264	0.3068	0.0258		
$\hat{\phi}_2$	−0.1613	0.0201	−0.2179	0.0182		
$\hat{\theta}_1$	0.0761	0.0248	0.1161	0.0256	0.0997	0.0071
$\hat{\theta}_2$	−0.1385	0.0156	−0.1970	0.0193	0.0703	0.0060
$\hat{\theta}_3$	−0.0280	0.0093	0.0038	0.0087	0.0588	0.0057
1249-Day Period: April 21, 1982–March 31, 1987						
$\bar{R}^2$	0.2553					
$\hat{\mu}$	0.0003	0.0002				
$\hat{\phi}_1$	0.5410	0.0187				
$\hat{\phi}_2$	−0.1269	0.0138				
$\hat{\theta}_1$	0.1248	0.0170				
$\hat{\theta}_2$	−0.1504	0.0100				
$\hat{\theta}_3$	−0.0534	0.0062				

[a] Parameter estimates obtained from times series regression across 5-minute returns during each trading day of the sample period. For the stock indexes, the first two returns each day are excluded. The average parameter estimate is computed across days.

[b] The standard error of the parameter estimate is computed across days.

In Table 2, the average parameter estimates for the fitted *MA*(3) and *ARMA*(2,3) models are reported. The IBM results show that the average *MA* coefficients are 0.0997, 0.0703, and 0.0588 for orders one through three, respectively. The coefficients have the expected sign and are significantly different from zero. The results for IBM in Table 3 show that after the effect of the bid/ask

[13] Infrequent trading does not induce positive serial correlation in the observed returns of individual stocks as it did for stock portfolio returns. The observed stock return may differ from the true return in a given time interval because the stock did not trade, however, the assumptions of independence and stationarity in the true generating process ensure that the observed return is an unbiased estimate of true return.

spread is removed, the serial correlation in the return innovations of IBM is less than 0.01 for all lags one through 12.

The Table 2 results for the stock indexes show that the coefficients of the *ARMA*(2,3) are generally significantly different from zero. The expected signs of the coefficients are unknown because of the simultaneous infrequent trading and bid/ask price effects at work on the observed return series. The average R^2 values indicate that infrequent trading and bid/ask price movements explain a good deal of the variation in observed index returns. In the overall 1,249-day sample period for the S&P 500 index, for example, the average R^2 exceeds 25 percent. The average R^2 is lower for the 609-day subperiod, reflecting the fact that there was considerably more trading activity in the stock market during the latter half of the overall sample period. The average R^2 for the MMI (0.0935) is lower than that of the S&P 500 (0.1747) during the same period, reflecting the fact that the narrowly based MMI has more active stocks.

As a precautionary measure, the residuals from the *ARMA*(2,3) regression were examined for violations of the assumption of homoskedasticity. Recall that in the development of the final *ARMA* specification, it was shown that the contemporaneous error, the lag one error, and the errors of higher order lags may have different variances. Using Engle's (1982) test for first-order *ARCH*, we found that the assumption of homoskedasticity was well supported empirically. The greatest proportion of rejections of the null hypothesis of constant variance was found for the S&P 500 index return regressions, and, even in this worst case, rejections were found in only 7 percent of the 1,249 days of the overall sample period.

The *ARMA* model appears to do very well at purging the effects of infrequent trading and bid/ask spreads, as evidenced by the serial correlation estimates reported in Table 3. The serial correlations are negligible at all lags. For example, none of the serial correlations for the S&P 500 index returns exceeds an absolute value of 0.0163 for the overall 1,249-day sample period. Compared with the serial correlation estimates in Table 1, the results are impressive. In addition, they provide reassurance that $\epsilon_{S,t}$ measures primarily true return innovation ($\eta_{S,t}$) in period t.

Another perspective on how well the *ARMA* model does at purging the effects of infrequent trading and bid/ask prices may be gathered from the multiple regression results reported in Table 4. When observed stock portfolio returns are regressed on lead, contemporaneous, and lag observed returns of other stock portfolios, the results indicate that IBM returns lead MMI returns, and MMI returns, in turn, lead S&P 500 returns. This is exactly the pattern that is expected since IBM trades more frequently than the stocks within the MMI, on average, and the stocks within the MMI trade more frequently than the stocks within the S&P 500, on average. When return innovations from the *ARMA* models are used, however, the estimated regressions for the S&P 500 and the MMI, the S&P 500 and IBM, and the MMI and IBM are remarkably similar. The strongest relation is at the contemporaneous level, the lead/lag one coefficients are approximately the same size, and higher order leads and lags are unimportant. For example, when raw returns are used, IBM appears to lead the S&P 500, on average. Both the lag one and lag two coefficient estimates $\hat{\beta}_1$, and $\hat{\beta}_2$ have large positive and

TABLE 3

Estimated Serial Correlation Coefficients of Return Innovations[a] of the S&P 500 Index (ϵ_S), the Major Market Index (ϵ_M), and the IBM (ϵ_I)

	S&P 500 Index (S) $\rho_k(\epsilon_{S,t},\epsilon_{S,t-k})$			MM Index (M) $\rho_k(\epsilon_{M,t},\epsilon_{M,t-k})$			IBM (I) $\rho_k(\epsilon_{I,t},\epsilon_{I,t-k})$		
Lag k	No. of Obs.	$\hat{\rho}_k^b$	$t(\hat{\rho}_k)^c$	No. of Obs.	$\hat{\rho}_k^b$	$t(\hat{\rho}_k)^c$	No. of Obs.	$\hat{\rho}_k^b$	$t(\hat{\rho}_k)^c$
609-Day Period: July 23, 1984–December 31, 1986									
1	41,975	0.0085	1.75	41,865	−0.0012	−0.25	43,978	0.0046	0.97
2	41,366	0.0052	1.06	41,256	0.0092	1.87	43,369	0.0061	1.28
3	40,757	0.0033	0.67	40,647	−0.0003	−0.07	42,760	0.0060	1.23
4	40,148	−0.0021	−0.42	40,038	0.0008	0.16	42,151	−0.0089	−1.82
5	39,539	0.0076	1.51	39,429	−0.0012	−0.23	41,542	0.0035	0.71
6	38,930	−0.0071	−1.40	38,820	−0.0064	−1.25	40,933	0.0015	0.31
7	38,321	−0.0088	−1.72	38,211	−0.0212	−4.14	40,324	−0.0011	−0.22
8	37,712	0.0070	1.37	37,602	−0.0015	−0.30	39,715	−0.0019	−0.38
9	37,103	−0.0056	−1.08	36,993	−0.0046	−0.88	39,106	0.0006	0.12
10	36,494	−0.0101	−1.93	36,384	−0.0038	−0.73	38,497	0.0046	0.91
11	35,885	−0.0115	−2.18	35,775	−0.0040	−0.76	37,888	−0.0096	−1.86
12	35,276	−0.0060	−1.12	35,166	0.0027	0.50	37,279	−0.0020	−0.38
1249-Day Period: April 21, 1982–March 31, 1987									
1	84,454	0.0071	2.06						
2	83,205	0.0053	1.52						
3	81,956	0.0068	1.95						
4	80,707	0.0050	1.41						
5	79,458	0.0052	1.48						
6	78,209	−0.0042	−1.18						
7	76,960	−0.0119	−3.30						
8	75,711	0.0017	0.46						
9	74,462	−0.0005	−0.15						
10	73,213	−0.0082	−2.22						
11	71,964	−0.0163	−4.37						
12	70,715	−0.0067	−1.77						

[a] The return innovations for the stock indexes are the residuals from an *ARMA*(2,3) model fit to the index return series each day, where the first two 5-minute returns are excluded. The IBM return innovations are the residuals from an *MA*(3) model fit to the IBM return series each day.

[b] The estimated lag k serial correlation coefficient across all 5-minute returns in all days of the period excluding overnight returns. For the stock indexes, the first two returns each day are excluded.

[c] The t-ratio corresponding to the null hypothesis ρ_k equals zero.

significant values, 0.1756 and 0.0944, respectively, and the lag three coefficient is approximately the same size as the lead one coefficient, 0.0461 and 0.0453, respectively. After the effects of infrequent trading and bid/ask spreads are removed, however, by far the dominant relation at the contemporaneous level, 0.1500, the lead/lag one coefficient estimates are approximately the same size, 0.0438 and 0.0686, respectively, and the higher order lead and lag coefficients have negligible values. In summary, when return innovations are considered, the S&P 500 and MM indexes do about as well predicting IBM as IBM does predicting the indexes. We now examine whether index futures returns can predict the return innovations of the indexes and IBM.

V. Do Stock Index Futures Returns Lead Stock Index Returns?

In perfectly efficient and continuous markets, the rates of return of the stock index and the index futures contract are perfectly positively and contemporane-

TABLE 4

Parameter Estimates from Regressions Using Stock Index and IBM Returns (R_t^o) and Return Innovations (ϵ_t)[a] during the 609-Day Sample Period July 23, 1984–Dec. 31, 1986 (Number of 5-minute return/return innovation observations common to all series is 38,475)

	S&P 500 Index (S) on MM Index (M)		S&P 500 Index (S) on IBM (I)		MM Index (M) on IBM (I)	
Parameter	Parameter Estimate[b]	t-Ratio[c]	Parameter Estimate[b]	t-Ratio[c]	Parameter Estimate[b]	t-Ratio[c]
Returns: $R^o_{i,t} = \alpha + \Sigma^3_{k=-3}\beta_k R^o_{j,t-k} + u_t$; $i = S, M; j = M, I; i \neq j$						
R^2	0.7105		0.4045		0.3472	
$\hat{\alpha}$	−0.0001	−0.55	0.0002	1.22	0.0004	1.25
$\hat{\beta}_{-3}$	−0.0074	−4.11	−0.0063	−4.01	−0.0129	−5.12
$\hat{\beta}_{-2}$	−0.0149	−7.91	−0.0055	−3.46	−0.0005	−0.21
$\hat{\beta}_{-1}$	0.0663	35.02	0.0453	28.33	0.0800	31.26
$\hat{\beta}_0$	0.4572	240.43	0.1881	117.09	0.3005	116.99
$\hat{\beta}_1$	0.1736	91.04	0.1756	108.49	0.2238	86.49
$\hat{\beta}_2$	0.0612	31.90	0.0944	58.14	0.0821	31.61
$\hat{\beta}_3$	0.0339	18.08	0.0461	28.54	0.0263	10.20
Return Innovations: $\epsilon_{i,t} = \alpha + \Sigma^3_{k=-3}\beta_k \epsilon_{j,t-k} + u_t$; $i = S, M; j = M, I; i \neq j$						
R^2	0.4522		0.2173		0.2243	
$\hat{\alpha}$	−0.0001	−0.84	−0.0002	−1.22	−0.0002	−0.66
$\hat{\beta}_{-3}$	−0.0058	−2.60	−0.0060	−3.73	−0.0137	−5.15
$\hat{\beta}_{-2}$	−0.0134	−5.94	−0.0041	−2.52	0.0015	0.55
$\hat{\beta}_{-1}$	0.0599	26.32	0.0438	26.78	0.0707	26.27
$\hat{\beta}_0$	0.3980	174.57	0.1500	91.59	0.2523	93.52
$\hat{\beta}_1$	0.0625	27.32	0.0686	41.56	0.1151	42.33
$\hat{\beta}_2$	0.0056	2.43	0.0056	3.37	0.0084	3.09
$\hat{\beta}_3$	−0.0021	−0.89	0.0043	2.60	0.0035	1.29

[a] The return innovations for the stock indexes are the residuals from an *ARMA*(2,3) model fit to the index return series each day, where the first two 5-minute returns are excluded. The IBM return innovations are the residuals from an *MA*(3) model fit to the IBM return series each day.

[b] Parameter estimates obtained from times series regression across all 5-minute returns in all days of the period excluding overnight returns. For the stock indexes, the first two returns each day are excluded.

[c] The t-ratio corresponding to the null hypothesis, the respective coefficient equals zero.

ously correlated; that is, the prices of the stock index and the futures simultaneously reflect new information as it reaches the marketplace. If there is price discovery in the futures market and/or time delays in reporting the index, however, the futures returns will tend to lead the stock index returns. In this section, a multiple regression framework is used to evaluate the extent to which index futures returns lead or lag stock returns.

Instruments for the stock returns ($Z_{S,t}$) are used as dependent variables in a regression on lead, contemporaneous, and lag futures returns as the independent variables;[14] that is,

$$Z_{S,t} = \alpha + \sum_{k=-3}^{3} \beta_k R^o_{F,t-k} + u_t \,. \tag{9}$$

[14] The regression framework used here is like that suggested by Sims (1972), except that a contemporaneous variable is also included as a regressor. Furthermore, only the results of the regression of futures returns on lag, contemporaneous, and lead stock index returns are reported. Qualitatively, the same results are obtained by regressing the futures returns on the lag, contemporaneous,

As was noted in Section IV, two types of instrumental variables are used: (a) return innovations generated by an *ARMA*(2,3) model fitted to index returns (i.e., $Z_{S,t} \equiv \epsilon_{S,t}$ for the S&P 500 index and $Z_{S,t} \equiv \epsilon_{M,t}$ for the MM index); and (b) return innovations generated by an *MA*(3) model fitted to IBM's returns (i.e., $Z_{S,t} \equiv \epsilon_{I,t}$). If the theoretical relation (2) is correct, the contemporaneous variable coefficient β_0 should be greater than zero in all three regressions, while all other coefficients should not be different from zero. Positive values for the coefficients at lags $k = 1, 2, 3$ would indicate that the returns in the futures market tend to lead those in the stock market, and positive values for the coefficients at leads $k = -1, -2, -3$ would indicate that the stock market tends to lead the futures market. For purposes of comparison, regression (9) is also performed using observed stock and stock index returns as the dependent variable.

Prior to investigating the results, it is worthwhile to highlight two aspects of the ordinary least squares (*OLS*) regression model (9). First, since the return innovations are pooled across days to act as the dependent variable, the regression assumption of homoskedasticity may be violated. Recall that the *ARMA* parameters were estimated each day to account for different levels of trading activity across days. For the same reason, the variance of the return innovations generated by the *ARMA* model may vary from day to day. Second, the observed return innovation of the stock portfolio used as the dependent variable in the regression model (9) is the sum of three components: (a) an error from the infrequent trading model, (b) an average bid/ask price error, and (c) a true return innovation in the stock portfolio. Of these, the important component is probably (c), which represents the stock portfolio's reaction to new information disseminating into the market. When the regression model (9) is estimated, however, the observed return innovation of the stock portfolio that appears on the left-hand side of the regression should be purged of the true return innovation by the futures returns that appear on the right-hand side (the futures contract is a derivative instrument whose price is determined by the same set of information as the underlying commodity; in this case, a stock index portfolio), so the error term in the regression model (μ_t) is left to pick up any lingering effects of misspecification of the *ARMA* model. While neither of these problems causes the *OLS* parameter estimates to be biased, the estimators are not efficient and the standard errors of the regression coefficients may be affected. For this reason, it is important to investigate the properties of μ_t after the regression model is estimated.

A. S&P 500 Futures Results

Regression results for the S&P 500 futures contract during the 1,249-day sample period are reported in the first two columns in Table 5. In the returns regression, the largest coefficient estimate, 0.2032, is for the lag one futures return, indicating that the price changes in the S&P 500 futures market lead the price changes of the underlying index, but this relation is illusory, given the infrequent trading of the stocks within the index. In the return innovations regression, the largest coefficient estimate, 0.1338, is for the contemporaneous vari-

and lead instrument returns. (Granger (1969) and Pierce and Haugh (1977) offer other approaches to the "causality" determination.)

able, as is expected in efficiently functioning markets; however, the markets are clearly not moving in perfect unison. The large, positive coefficient of the lag one futures return, 0.1015, together with its *t*-ratio, 85.72, indicate that the futures return last period is also strongly associated with the current index return. Moreover, the lag two coefficient, $\hat{\beta}_2$, is also positive, albeit small in its relative magnitude, and less significant. In addition, the R^2 value of 0.2132 indicates that the presence of the lag and lead futures index returns increases considerably the explained variation of the index returns from the case in which only the contemporaneous futures return is used. The R^2 value in the simple linear regression of the return innovation of the S&P 500 index on the index futures return is 0.1308.

The lead one futures return coefficient, 0.0194, reveals the presence of feedback in the index return/futures return relation. The effect is small when viewed relative to the lead of futures returns, but, nonetheless, it supports the notion that occasionally the index leads the futures. The overall implication of the results is that the S&P 500 futures returns have a significantly greater tendency to lead than to lag the return innovations of the S&P 500 index.

In the introduction of this section, we noted that the interpretation of the *OLS* regression results may be influenced by possible heteroskedasticity and serial dependence in the error term. Heteroskedasticity, if it appears, would likely be driven by the nonconstant variance in the dependent variable of (9), which was generated by an *ARMA* regression on 5-minute returns each trading day. Since we argued that daily estimation of the *ARMA* model was necessary due to different levels of trading activity, it is also reasonable to argue that the variance of the residual from the *ARMA* estimation varies from day to day. To test the impact of this argument, weighted least squares regressions (*WLS*) were performed.[15] The *WLS* results are remarkably similar to the *OLS* results reported in Table 5. For example, the coefficient of the contemporaneous futures return is estimated to be 0.1383 and has a *t*-ratio of 105.64 in the *WLS* regression, where the values were 0.1338 and 113.50, respectively, under *OLS* regression. The lag one futures return coefficient estimate increased from 0.1015 under *OLS* to 0.1384 under *WLS*, and the corresponding *t*-ratios are 85.72 and 104.87, respectively. The lead one futures return coefficient and its *t*-ratio also increased slightly in magnitude—from 0.0194 and 16.54, respectively, in the *OLS* regression, to 0.0280 and 21.52, respectively, in the *WLS* regression. Clearly, accounting for heteroskedasticity does not alter the interpretation of the lead/lag structure. Under both estimation procedures, the lag one futures return is large, positive, and highly significant, while the lead one futures return coefficient is small, positive, and significant. The interpretation remains that the S&P 500 futures returns tend to lead rather than lag the stock market returns, although the effect is not completely unidirectional.

Serial dependence in the error term of (9) also may affect the interpretation of the regression results. Upon examining the residuals of the *OLS* regression of the observed S&P 500 index return innovations on lead, contemporaneous, and

[15] All variables (returns) for each day in (9) were deflated by the standard error of the estimate from the *ARMA* regression for that day. Naturally, the *WLS* procedure introduced a new independent variable—one over the standard error of the estimate—and the need to suppress the intercept in the regression.

TABLE 5

Parameter Estimates from Regressions of Stock Index Returns (R^o_i) and Return Innovations (ϵ_i)[a] on Lag, Contemporaneous, and Lead Nearby Futures S&P 500 Returns ($R^o_{F,t}$)

	1,249-Period April 21, 1982–March 31, 1987 S&P 500 Index (S)		609-Period July 23, 1984–Dec. 31, 1986 S&P 500 Index (S)		609-Period July 23, 1984–Dec. 31, 1986 MM Index (M)		609-Period July 23, 1984–Dec. 31, 1986 IBM (I)	
Parameter	Parameter Estimate[b]	t-Ratio[c]	Parameter Estimate[b]	t-Ratio[c]	Parameter Estimate[b]	t-Ratio[c]	Parameter Estimate[b]	t-Ratio[c]
Returns: $R^o_{i,t} = \alpha + \Sigma^3_{k=-3}\beta_k R^o_{F,t-k} + u_t$; $i = S, M, I$								
No. of Obs.	78,209		38,930		38,820		40,933	
R^2	0.4730		0.5188		0.4399		0.2440	
$\hat{\alpha}$	−0.0001	−1.08	0.0001	0.46	0.0002	0.60	−0.0004	−0.86
$\hat{\beta}_{-3}$	−0.0077	−6.57	−0.0038	−2.03	−0.0073	−2.34	−0.0117	−2.05
$\hat{\beta}_{-2}$	−0.0158	−13.48	−0.0192	−10.21	−0.0209	−6.71	−0.0122	−2.13
$\hat{\beta}_{-1}$	0.0213	18.10	0.0195	10.34	0.0335	10.74	0.0872	15.19
$\hat{\beta}_0$	0.1690	142.93	0.2210	116.37	0.3706	118.40	0.5659	98.11
$\hat{\beta}_1$	0.2032	171.14	0.2734	143.52	0.3768	119.87	0.3340	57.89
$\hat{\beta}_2$	0.1330	111.45	0.1541	80.61	0.1325	42.05	0.0142	2.46
$\hat{\beta}_3$	0.0798	66.50	0.0710	37.00	0.0361	11.41	−0.0151	−2.61
Return Innovations: $\epsilon_{i,t} = \alpha + \Sigma^3_{k=-3}\beta_k R^o_{F,t-k} + u_t$; $i = S, M, I; i \neq j$								
No. of Obs.	78,209		38,930		38,820		40,933	
R^2	0.2132		0.2730		0.2802		0.2393	
$\hat{\alpha}$	−0.0002	−1.73	−0.0003	−1.55	−0.0003	−1.01	−0.0004	−0.74
$\hat{\beta}_{-3}$	−0.0094	−8.04	−0.0134	−6.75	−0.0220	−6.68	−0.0291	−5.27
$\hat{\beta}_{-2}$	−0.0153	−13.04	−0.0228	−11.43	−0.0326	−9.89	−0.0351	−6.36
$\hat{\beta}_{-1}$	0.0194	16.54	0.0167	8.38	0.0212	6.43	0.0652	11.79
$\hat{\beta}_0$	0.1338	113.50	0.1833	91.23	0.3130	94.38	0.5206	93.67
$\hat{\beta}_1$	0.1015	85.72	0.1521	75.50	0.2514	75.49	0.3458	62.20
$\hat{\beta}_2$	0.0153	12.87	0.0301	14.88	0.0392	11.74	0.0638	11.45
$\hat{\beta}_3$	0.0059	4.92	0.0148	7.29	0.0264	7.87	0.0256	4.60

[a] The return innovations for the stock indexes are the residuals from an *ARMA*(2,3) model fit to the index return series each day, where the first two 5-minute returns are excluded. The IBM return innovations are the residuals from an *MA*(3) model fit to the IBM return series each day.

[b] Parameter estimates obtained from times series regression across all 5-minute returns in all days of the period excluding overnight returns. For the stock indexes, the first two returns each day are excluded.

[c] The t-ratio corresponding to the null hypothesis, the respective coefficient equals zero.

lag futures returns, negligible serial correlation was found at all lags except lag one. The first-order serial correlation in the residuals was estimated to be −0.1595 and is significantly different from zero. While this result may seem surprising, given that virtually no serial dependence exists in either the dependent or independent variables in (9), recall the components of the dependent variable. One component is the true return innovation of index portfolio, which is now being controlled for by the presence of index futures returns on the right-hand side of the regression. The other components are the infrequent trading and bid/ask price errors, both of which may induce negative serial dependence in the residuals of (9). The *ARMA* model fitted to the observed portfolio returns each day is an average of the infrequent trading and bid/ask price effects for the day; that is, it assumes that trading frequency is uniform throughout the day and that last transaction prices for the stocks in the portfolio are about evenly split between bid prices and ask prices. Now, suppose that during the trading day, some unexpected favorable market news is announced, and all stocks trade immediately in reaction to the news. The error term in the *ARMA* regression will be inordinately large for two reasons: (a) the true return innovation in the market is high, and (b) the *ARMA* model is misspecified. The former effect is captured by the futures returns in the regression (9). The latter effect is captured in the residual term μ_t. The *ARMA* model misspecification in this illustration arises from two sources: (a) all stocks trade in the interval that the news arrives, while they do not on average, and (b) all stocks in the portfolio trade at the ask prices where typically the stocks are evenly balanced between bids and asks. In the next interval, the error will tend to be smaller or negative since the mean error across the day is equal to zero by construction. The reason that no serial dependence appears in the residuals of (8), $\epsilon_{S,t}$, while negative serial dependence appears in the residuals of (9), μ_t, is that the dominant proportion of the variation of $\epsilon_{S,t}$ is the true return innovation of the index portfolio. With that effect controlled for by the futures returns in (9), the infrequent trading and bid/ask price effects are more free to appear.

To account for the negative serial dependence in the residuals of (9), generalized least squares regression was performed. The estimated coefficients of the lag one, contemporaneous, and lead one futures returns were 0.1011, 0.1321, and 0.0191, with *t*-ratios of 84.26, 110.80, and 16.11, respectively. As was the case for heteroskedasticity, the more complex estimation procedure does not alter the interpretation of the results. For this reason, all subsequent interpretations are based on the simpler *OLS* regression results.

The S&P 500 results for the 609-day subperiod indicate that the relation between returns in the stock and futures markets has grown tighter. The magnitudes of the contemporaneous and lag one coefficients are larger than in the 1,249-day sample period. In addition, the R^2 is larger, indicating that the correlation between returns in the two markets is higher. (Market maturation effects are examined in more detail later in this section.) These results also can be contrasted with those in Table 4 to reaffirm the ability of the *ARMA* model in removing the effects of infrequent trading and the interpretation that the futures market tends to lead the stock market. Where the *ARMA* model produced a return innovation series for the S&P 500 that was contemporaneous with the return innovation series of MMI and IBM on average (as shown by the large contemporaneous coefficient and the similar magnitudes of the lead/lag one coefficients in Table 4), the same return innovations for the S&P 500 index lag the S&P 500 futures returns as reported in Table 5. While the contemporaneous coefficient estimate is largest, the lag one coefficient is also large while the lead one coefficient is negligible.

The final four columns in the bottom panel of Table 5 contain the results of regressions using MMI and IBM return innovations as dependent variables. Although different instruments for the true returns of the S&P 500 index are being used, the results are qualitatively the same. The levels of the regression coefficients cannot be compared meaningfully since different purging regressions are used for each instrument in each sample period. The *t*-ratios (levels of significance) can be compared, however, and the dominant effect is at the contemporaneous level, with the futures return at lag one also having a strong positive effect. In the case of IBM, these results are particularly important since IBM is less contaminated by the effects of infrequent trading and of possible misspecification in the *ARMA* infrequent trading model. While Table 4 shows that even IBM, considered to be the "bellwether" stock, does not lead the return innovations of the S&P 500 and MM indexes on average, Table 5 shows that index futures do lead index return innovations on average. Taken together, the results of Table 5 imply that index futures prices respond more rapidly to economic events than stock prices.

Considering that the computation and dissemination of the S&P 500 and MM index levels are handled by two entirely independent reporting services,[16] the fact that the coefficient structure in the S&P 500 and MMI regressions is so similar is somewhat surprising. The evidence suggests that, while there may be (a) a delay in recording the stock transaction information on the floor of the NYSE,[17] no significant delays occur in either (b) the computation and transmission of new index values given new stock transaction information, or (c) the recording of the new stock index levels at the futures exchange.

B. MMI Futures Results

The generality of the S&P 500 futures results can be evaluated using data from another stock index futures market—in this case, the CBOT's Major Market Index futures contract. Table 6 contains the results of regressions in which stock returns and stock return innovations are regressed on the lead, contemporaneous, and lag returns of the MMI futures contract. In the overall 678-day period, the contemporaneous and lag one futures returns in the return innovations regression have coefficients 0.3479 and 0.2533. This result is consistent with the S&P 500 result in the sense that the dominant relation is contemporaneous with a strong tendency for the futures returns to lead the stock index returns by about five minutes. Note that the coefficient structure in the return regression (top panel) is not as dramatically different from the return innovation regression (bottom panel) as it was in the S&P 500 results (see Table 5). This simply reflects the fact that the MMI stocks have much more active secondary markets on average than do the stocks within the S&P 500 index.

The MMI results are different from the S&P 500 results in some subtle ways. For example, during the 678-day sample period, the lead one coefficient estimate of −0.0064 indicates that, unlike the S&P 500 case, there is little, if any, feedback from the stock market to the MM futures market.

C. Market Maturation Effects

The results reported in Tables 5 and 6 make only a casual attempt to examine differences in the temporal relation between futures and index returns as the index futures markets have matured. To examine market maturation effects more

[16] Recall that the S&P 500 index levels are computed and reported by ADP Brokerage Information Services, while the MMI levels are computed and reported by the AMEX.

[17] Conversations with specialists, however, indicate that even in very active markets, the recording delay never exceeds a couple of minutes.

TABLE 6

Parameter Estimates from Regressions of Stock Index Returns (R_t^o) and Return Innovations (ϵ_t)[a] on Lag, Contemporaneous, and Lead Nearby MMI Futures Returns ($R_{F,t}^o$)

	678-Period July 23, 1984– March 31, 1987 MM Index (*M*)		609-Period July 23, 1984– Dec. 31, 1986 MM Index (*M*)		609-Period July 23, 1984– Dec. 31, 1986 S&P 500 Index (*S*)		609-Period July 23, 1984– Dec. 31, 1986 IBM (*I*)	
Parameter	Parameter Estimate[b]	*t*-Ratio[c]	Parameter Estimate[b]	*t*-Ratio[c]	Parameter Estimate[b]	*t*-Ratio[c]	Parameter Estimate[b]	*t*-Ratio[c]
Returns: $R_{i,t}^o = \alpha + \Sigma_{k=-3}^{3} \beta_k R_{F,t-k}^o + u_t$; $i = S, M, I$								
No. of Obs.	43,217		38,681		38,543		40,642	
R^2	0.4374		0.4387		0.4663		0.2186	
$\hat{\alpha}$	−0.0001	−0.53	−0.0002	−0.61	−0.0001	−0.76	−0.0006	−1.17
$\hat{\beta}_{-3}$	−0.0041	−1.36	−0.0067	−2.16	−0.0009	−0.45	−0.0079	−1.35
$\hat{\beta}_{-2}$	−0.0148	−4.89	−0.0143	−4.57	−0.0140	−7.02	−0.0107	−1.82
$\hat{\beta}_{-1}$	0.0058	1.91	0.0105	3.36	0.0098	4.89	0.0787	13.33
$\hat{\beta}_0$	0.4023	131.32	0.3835	121.56	0.2125	105.38	0.5363	90.33
$\hat{\beta}_1$	0.3768	122.30	0.3713	116.73	0.2595	127.66	0.3390	57.05
$\hat{\beta}_2$	0.1459	47.20	0.1564	48.97	0.1581	77.48	0.0451	7.56
$\hat{\beta}_3$	0.0522	16.83	0.0525	16.38	0.0808	39.42	−0.0072	−1.20
Return Innovations: $\epsilon_{i,t} = \alpha + \Sigma_{k=-3}^{3} \beta_k R_{F,t-k}^o + u_t$; $i = S, M, I$								
No. of Obs.	43,217		38,681		38,543		40,642	
R^2	0.2937		0.2845		0.2235		0.2170	
$\hat{\alpha}$	−0.0007	−2.58	−0.0005	−1.57	−0.0001	−0.44	−0.0005	−1.07
$\hat{\beta}_{-3}$	−0.0203	−6.39	−0.0214	−6.53	−0.0096	−4.42	−0.0249	−4.42
$\hat{\beta}_{-2}$	−0.0251	−7.87	−0.0236	−7.18	−0.0200	−9.24	−0.0297	−5.26
$\hat{\beta}_{-1}$	−0.0064	−2.01	−0.0015	−0.45	0.0044	2.01	0.0567	10.00
$\hat{\beta}_0$	0.3479	108.11	0.3310	99.36	0.1804	82.23	0.4941	86.70
$\hat{\beta}_1$	0.2533	78.28	0.2434	72.50	0.1412	63.81	0.3478	60.98
$\hat{\beta}_2$	0.0601	18.52	0.0624	18.52	0.0405	18.22	0.0900	15.71
$\hat{\beta}_3$	0.0352	10.81	0.0315	9.29	0.0216	9.66	0.0325	5.65

[a] The return innovations for the stock indexes are the residuals from an *ARMA*(2,3) model fit to the index return series each day, where the first two 5-minute returns are excluded. The IBM return innovations are the residuals from an *MA*(3) model fit to the IBM return series each day.

[b] Parameter estimates obtained from times series regression across all 5-minute returns in all days of the period excluding overnight returns. For the stock indexes, the first two returns each day are excluded.

[c] The *t*-ratio corresponding to the null hypothesis, the respective coefficient equals zero.

directly, the regression model (7) is estimated for each of four subperiods of the 1,249-day S&P 500 sample and for each of four subperiods of the 678-day MMI sample. The S&P 500 results are reported in Table 7, and the MMI results in Table 8.

The S&P 500 results in Table 7 show that the temporal relation between futures returns and index returns has changed through time. In the first subperiod, for example, the contemporaneous and lag futures return coefficients are generally smaller than in the other subperiods. Moreover, the lead one coefficient estimate $\hat{\beta}_{-1}$ is more significant than in the other subperiods, indicating that there was a strong tendency of index returns to lead futures returns early in the life of the S&P 500 futures market. Perhaps market makers in the S&P 500 futures pit were revising futures prices as new index information was being obtained. None of the subperiod results shows any tendency for the futures prices to

TABLE 7

Parameter Estimates from Regressions of Stock Index Returns (R_t^o) and Return Innovations (ϵ_t)[a] on Lag, Contemporaneous, and Lead Nearby S&P 500 Futures Returns ($R_{F,t}^o$)

Parameter	312-Period April 21, 1982– July 13, 1983 Parameter Estimate[b]	t-Ratio[c]	312-Period July 14, 1983– Oct. 8, 1984 Parameter Estimate[b]	t-Ratio[c]	312-Period Oct. 9, 1984– Jan. 3, 1986 Parameter Estimate[b]	t-Ratio[c]	313-Period Jan. 6, 1986– March 31, 1987 Parameter Estimate[b]	t-Ratio[c]
Returns: $R_{S,t}^o = \alpha + \Sigma_{k=-3}^{3} \beta_k R_{F,t-k}^o + u_t$								
No. of Obs.	18,920		18,966		19,383		20,940	
R^2	0.5209		0.5310		0.4913		0.5564	
$\hat{\alpha}$	−0.0004	−1.57	−0.0004	−2.18	0.0002	1.16	−0.0001	−0.21
$\hat{\beta}_{-3}$	−0.0112	−6.71	−0.0047	−2.11	−0.0068	−2.67	−0.0027	−1.03
$\hat{\beta}_{-2}$	−0.0061	−3.62	−0.0039	−1.75	−0.0113	−4.45	−0.0249	−9.43
$\hat{\beta}_{-1}$	0.0293	17.44	0.0280	12.61	0.0297	11.65	0.0194	7.32
$\hat{\beta}_0$	0.1088	64.47	0.1790	80.33	0.2118	82.26	0.2424	91.25
$\hat{\beta}_1$	0.1192	70.28	0.1829	81.79	0.2252	87.22	0.3341	125.54
$\hat{\beta}_2$	0.1151	67.39	0.1617	72.05	0.1594	61.48	0.1433	53.76
$\hat{\beta}_3$	0.0955	55.57	0.1158	51.20	0.0931	35.73	0.0433	16.21
Return Innovations: $\epsilon_{S,t} = \alpha + \Sigma_{k=-3}^{3} \beta_k R_{F,t-k}^o + u_t$								
No. of Obs.	18,920		18,966		19,383		20,940	
R^2	0.1797		0.2439		0.2328		0.3234	
$\hat{\alpha}$	−0.0003	−1.34	0.0002	0.81	−0.0002	−1.18	−0.0006	−1.83
$\hat{\beta}_{-3}$	−0.0050	−3.19	−0.0078	−3.67	−0.0165	−6.09	−0.0126	−4.46
$\hat{\beta}_{-2}$	−0.0023	−1.48	−0.0098	−4.60	−0.0158	−5.80	−0.0272	−9.61
$\hat{\beta}_{-1}$	0.0273	17.41	0.0175	8.18	0.0210	7.70	0.0188	6.65
$\hat{\beta}_0$	0.0801	50.94	0.1450	67.58	0.1718	62.52	0.1970	69.34
$\hat{\beta}_1$	0.0419	26.52	0.0729	33.86	0.1122	40.70	0.2076	72.91
$\hat{\beta}_2$	0.0141	8.86	0.0322	14.89	0.0412	14.89	0.0111	3.90
$\hat{\beta}_3$	0.0013	0.81	0.0071	3.27	0.0236	8.48	0.0138	4.83

[a] The return innovations for the stock indexes are the residuals from an *ARMA*(2,3) model fit to the index return series each day, where the first two 5-minute returns are excluded.

[b] Parameter estimates obtained from times series regression across all 5-minute returns in all days of the period excluding overnight returns and the first two returns each trading day.

[c] The *t*-ratio corresponding to the null hypothesis, the respective coefficient equals zero.

overshoot their equilibrium values and then fall back into alignment with respect to the stock index level. Such a phenomenon would be indicated if the coefficients of the lead futures returns were significantly large negative values.

Other indications that the S&P 500 futures market matured during the five-year sample period are also present. For example, the R^2 of the regression (9) increases from 0.1797 in the first subperiod to 0.3234 in the last. Clearly the comovement of intraday returns in the two markets became stronger through time. Another indication is that the number of meaningful coefficients narrows to only two in the last subperiod from as many as four in the earlier subperiods. On the basis of these results, it seems fair to conclude that the S&P 500 futures and stock markets have become more closely integrated through time, perhaps, as a result of more active index arbitrage and/or more efficient program trading. Nevertheless, it is surprising that the proportion of the total variation in stock returns explained by futures returns is less than 35 percent in all subperiods. Apparently,

TABLE 8

Parameter Estimates from Regressions of Stock Index Returns (R_t^o) and Return Innovations (ϵ_t)[a] on Lag, Contemporaneous, and Lead Nearby MMI Futures Returns ($R_{F,t}^o$)

	169-Period July 23, 1984–March 25, 1985		169-Period March 26, 1985–Nov. 22, 1985		170-Period Nov. 25, 1985–July 26, 1986		170-Period July 30, 1986–March 31, 1987	
Parameter	Parameter Estimate[b]	*t*-Ratio[c]	Parameter Estimate[b]	*t*-Ratio[c]	Parameter Estimate[b]	*t*-Ratio[c]	Parameter Estimate[b]	*t*-Ratio[c]
Returns: $R_{M,t}^o = \alpha + \Sigma_{k=-3}^{3} \beta_k R_{F,t-k}^o + u_t$								
No. of Obs.	10,205		10,422		11,007		11,250	
R^2	0.4513		0.3265		0.5364		0.4242	
$\hat{\alpha}$	0.0001	0.12	−0.0003	−0.64	−0.0003	−0.58	0.0001	−0.18
$\hat{\beta}_{-3}$	−0.0094	−1.92	−0.0006	−0.09	−0.0127	−2.17	0.0057	0.85
$\hat{\beta}_{-2}$	−0.0078	−1.60	−0.0031	−0.46	−0.0153	−2.59	−0.0207	−3.05
$\hat{\beta}_{-1}$	0.0280	5.69	0.0231	3.46	−0.0028	−0.48	−0.0224	−3.29
$\hat{\beta}_0$	0.2862	57.83	0.3499	52.14	0.4846	80.96	0.4549	67.16
$\hat{\beta}_1$	0.3094	62.06	0.3146	46.71	0.4286	70.79	0.4109	60.69
$\hat{\beta}_2$	0.1980	39.57	0.1902	28.27	0.1079	17.74	0.1136	16.74
$\hat{\beta}_3$	0.1002	20.04	0.0827	12.22	0.0029	0.47	0.0433	6.35
Return Innovations: $\epsilon_{M,t} = \alpha + \Sigma_{k=-3}^{3} \beta_k R_{F,t-k}^o + u_t$								
No. of Obs.	10,205		10,422		11,007		11,250	
R^2	0.2523		0.2180		0.3796		0.2996	
$\hat{\alpha}$	−0.0001	−0.10	−0.0005	−1.32	−0.0011	−1.75	−0.0017	−2.55
$\hat{\beta}_{-3}$	−0.0182	−3.55	−0.0168	−2.52	−0.0301	−4.79	−0.0160	−2.22
$\hat{\beta}_{-2}$	−0.0166	−3.22	−0.0162	−2.41	−0.0286	−4.51	−0.0319	−4.44
$\hat{\beta}_{-1}$	0.0118	2.28	0.0060	0.88	−0.0127	−1.98	−0.0337	−4.66
$\hat{\beta}_0$	0.2364	45.48	0.2945	43.58	0.4330	67.11	0.3973	55.19
$\hat{\beta}_1$	0.1925	36.76	0.2226	32.82	0.2735	41.90	0.2982	41.44
$\hat{\beta}_2$	0.0730	13.89	0.1059	15.62	0.0278	4.25	0.0596	8.27
$\hat{\beta}_3$	0.0310	5.91	0.0392	5.75	0.0277	4.21	0.0484	6.68

[a] The return innovations for the stock indexes are the residuals from an *ARMA*(2,3) model fit to the index return series each day, where the first two 5-minute returns are excluded.

[b] Parameter estimates obtained from times series regression across all 5-minute returns in all days of the period excluding overnight returns and the first two returns each trading day.

[c] The *t*-ratio corresponding to the null hypothesis, the respective coefficient equals zero.

noise from nonsynchronous trading within the 5-minute intervals as well as trading frictions limit the observed correlation between the returns in the two markets. A longer time interval between price observations would serve to increase this correlation.

In the last subperiod, the contemporaneous and lag one coefficient estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ are substantially higher and the other coefficients lower than in the other subperiods. This result is again likely attributable to the greater market activity during this subperiod. The average number of trades per day on the NYSE in the last subperiod was 78,197, compared with 58,168, 52,293, and 56,808 in subperiods 1 through 3, respectively.

The MMI results in Table 8 show a similar pattern. Although the overall time period is considerably shorter for the MM sample (678 days versus 1,249 days for the S&P 500), the estimated coefficients of the contemporaneous and lag one futures returns, $\hat{\beta}_0$ and $\hat{\beta}_1$ grow larger through time. In addition, the R^2 values have generally increased.

VI. Summary and Conclusions

This study investigates the time series properties of intraday returns of stock index and stock index futures contracts and finds several important results. First, S&P 500 and MM index futures returns lead stock index returns by about five minutes on average, but occasionally as long as ten minutes or more, after the observed stock index returns have been purged of infrequent trading and bid/ask price effects. Second, S&P 500 and MMI futures returns also tend to lead even the returns of actively traded stocks such as IBM. Third, although futures returns tend to lead stock returns, the effect is not completely unidirectional. There is a weak positive predictive effect of lag stock index returns on current futures returns; however, the effect has grown smaller as the futures markets have matured. Finally, the effects of infrequent trading and bid/ask price effects on observed, 5-minute rate of return series of the S&P 500 and MM stock indexes appear to be adequately described by an $ARMA(2,3)$ process. The observed, 5-minute return series of IBM follows an $MA(3)$ process.

In summary, the relations between the rates of return of the S&P 500 index and the S&P 500 index futures contracts and of the MMI stock index and MMI stock index futures contracts are as one might expect. The returns in the futures market and the stock market appear to be, in large part, contemporaneous. There is evidence that the futures market leads the stock market, and this is attributable, only in part, to the fact that not all stocks in the index trade continuously. The remaining predictive power of futures returns is evidence supporting the price discovery hypothesis that new market information disseminates in the futures market before the stock market, with index arbitrageurs then stepping in quickly to bring the cost-of-carry relation back into alignment.

Appendix: Derivation of a Model of Observed Portfolio Returns when Component Stocks Trade Infrequently

In this appendix, a model of stock portfolio returns that accounts for the effects of infrequent trading is developed. In the model, observed portfolio returns are expressed as a weighted average of contemporaneous and lagged true returns plus a mean zero error term, which is assumed to be independent and identically distributed through time; that is,

$$(1) \qquad R_t^o = \sum_{k=0}^{l} \omega_k R_{t-k} + \nu_t ,$$

where $l \equiv n-1$ and n is defined in the text. This equation may be restated as

$$(2) \qquad \begin{aligned} R_t^o &= \omega_0 \left(R_t + \frac{\omega_1}{\omega_0} R_{t-1} + \frac{\omega_2}{\omega_0} R_{t-2} + \cdots + \frac{\omega_l}{\omega_0} R_{t-l} \right) + \nu_t \\ &= \omega_0 \left(1 + \frac{\omega_1}{\omega_0} B + \frac{\omega_2}{\omega_0} B^2 + \cdots + \frac{\omega_l}{\omega_0} B^l \right) R_t + \nu_t \end{aligned}$$

for clarity. Dividing (2) by $(1+(\omega_1/\omega_0)B+(\omega_2/\omega_0)B^2+\ldots+(\omega_l/\omega_0)B^l)$ yields

$$\frac{R_t^o}{\left(1+\frac{\omega_1}{\omega_0}B+\frac{\omega_2}{\omega_0}B^2+\cdots+\frac{\omega_l}{\omega_0}B^l\right)}=\omega_0 R_t+\frac{1}{\left(1+\frac{\omega_1}{\omega_0}B+\frac{\omega_2}{\omega_0}B^2+\cdots+\frac{\omega_l}{\omega_0}B^l\right)}\nu_t\,. \tag{3}$$

Recalling the power series expansion

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^l+\cdots(|x|<1)\,,$$

the operator, $(1+(\omega_1/\omega_0)B+(\omega_2/\omega_0)B^2+\ldots+(\omega_l/\omega_0)B^l)^{-1}$, can be rewritten as

$$1-\left(\frac{\omega_1}{\omega_0}B+\cdots+\frac{\omega_l}{\omega_0}B^l\right)+\left(\frac{\omega_1}{\omega_0}B+\cdots+\frac{\omega_l}{\omega_0}B^l\right)^2-\cdots=1-\sum_{k=1}^{\infty}\gamma_k B^k\,, \tag{4}$$

where the γ_k's are functions of the ω_k's in (1). In general, the values of the γ_k's may be positive or negative but decline in absolute value as k increases (due to the assumption that $\omega_0>\omega_1>\ldots>\omega_l$). Substituting (4) into (3), we get,

$$R_t^o-\sum_{k=1}^{\infty}\gamma_k R_{t-k}^o=\omega_0 R_t+\nu_t-\sum_{k=1}^{\infty}\gamma_k \nu_{t-k}\,, \tag{5}$$

which can be rewritten as

$$R_t^o=\omega_0 R_t+\sum_{k=1}^{\infty}\gamma_k R_{t-k}^o+\nu_t-\sum_{k=1}^{\infty}\gamma_k \nu_{t-k}\,. \tag{6}$$

Equation (10) shows that observed stock portfolio returns follow an *ARMA(p,q)* process with a random intercept, where the coefficients on the lagged autoregressive and moving average parameters are the same.

References

Brady Commission Report. *Report of the Presidential Task Force on Market Mechanisms* (Jan. 1988).

CFTC Report. *Final Report on Stock Index Futures and Cash Market Activity during October 1987 to the U.S. Commodity Futures Trading Commission*. Prepared by the Division of Economic Analysis and the Division of Trading and Markets (Jan. 1988).

Cahn, K.; K. C. Chan; and G. A. Karolyi. "Transmissions of Volatility between Stock Index and Stock Index Futures Markets." Working Paper, Faculty of Finance, Ohio State University (1990).

Cohen, K. J.; S. A. Maier; R. A. Schwartz; and D. K. Whitcomb. "On the Existence of Serial Correlation in an Efficient Securities Market." *TIMS Studies in the Management Sciences*, 11 (1979), 151–168.

________________. *The Microstructure of Securities Markets*. Englewood Cliffs, NJ: Prentice-Hall, Inc. (1986).

Dimson, E. "Risk Measurement when Shares are Subject to Infrequent Trading." *Journal of Financial Economics*, 7 (June 1979), 197–226.

Engle, R. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation." *Econometrica*, 50 (July 1982), 987–1008.

Fama, E. F. *Foundations of Finance*, New York: Basic Books, Inc. (1976).

Fisher, L. "Some New Stock Market Indexes." *Journal of Business*, 39 (Jan. 1966), 191–225.

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

Harris, L. "The October 1987 S&P 500 Stock-Futures Basis." *Journal of Finance*, 44 (March 1989), 77–99.

Kawaller, I.; P. Koch; and T. Koch. "The Temporal Relationship between S&P 500 Futures Prices and the S&P 500 Index." *Journal of Finance*, 42 (Dec. 1987a), 1309–1329.

________________. "The Extent of Feedback between S&P 500 Futures Prices and the S&P 500 Index." Working Paper, Chicago Mercantile Exchange (Sept. 1987b).

Lo, A. W., and A. C. MacKinlay. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test." *Review of Financial Studies*, 1 (Spring 1988), 41–66.

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

Muthuswamy, J. "Nonsynchronous Trading and the Index Autocorrelation Problem." Working Paper, Graduate School of Business, Univ. of Chicago (1989).

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

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

Roll, R. "A Simple Implicit Measure of the Effective Bid/Ask Spread." *Journal of Finance*, 39 (Sept. 1984), 1127–1139.

SEC Report. *The October Market Break*. A report by the Division of Market Regulation, U.S. Securities and Exchange Commisson (Feb. 1988).

Scholes, M., and J. Williams. "Estimating Betas from Nonsynchronous Data." *Journal of Financial Economics*, 5 (Dec. 1977), 309–327.

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

Stephan, J. A., and R. E. Whaley. "Intraday Price Change and Trading Volume Relations in the Stock and Stock Option Markets." *Journal of Finance*, 45 (March 1990), 191–200.

Stoll, H. R. "Inferring the Components of the Bid/Ask Spread: Theory and Empirical Tests." *Journal of Finance*, 44 (March 1989), 115–134.

Stoll, H. R., and R. E. Whaley. "Transaction Costs and the Small Firm Effect." *Journal of Financial Economics*, 12 (June 1983), 57–79.

________________. *Expiration Day Effects of Index Options and Futures*. Monograph Series in Finance and Economics, New York Univ. Monograph 1986–3 (1986).

________________. "Program Trading and the Expiration Day Effects of Index Options and Futures." *Financial Analysts Journal* (March/April 1987), 16–28.

________________. "Stock Market Structure and Volatility." *Review of Financial Studies*, 3 (Spring 1990), 37–71.