THE JOURNAL OF FINANCE • VOL. XLIV, NO. 1 • MARCH 1989

The October 1987 S&P 500 Stock-Futures Basis

LAWRENCE HARRIS*

ABSTRACT

Five-minute changes in the S&P 500 index and futures contract are examined over a ten-day period surrounding the October 1987 stock market crash. Since nonsynchronous trading problems are severe in these data, new index estimators are derived and used. The estimators use the complete transaction history of all 500 stocks. Nonsynchronous trading explains part of the large absolute futures-cash basis observed during the crash. The remainder may be due to disintegration of the two markets. Even after adjustment for nonsynchronous trading, the index displays more autocorrelation than does the futures and the futures leads the index.

THE RELATION BETWEEN THE S&P 500 index futures market and the underlying cash market always has been of interest to investors, academics, and regulators. Recently, it became the focus of much new attention when volatility in these markets increased. One aspect of the relation between these two markets is the basis spread, the difference between the value of the underlying cash index and the futures price. The basis is studied because it is a key determinant of whether arbitrage opportunities exist, because variance in the basis is a measure of how well integrated the two markets are, and because the basis is related to tests for causality among the prices in the two markets.

The basis is normally calculated by subtracting the futures price from the most recent value of the S&P 500 index. The result, however, is unlikely to be the best measure of the true price relation between the two markets because the index often is not the best measure of the current value of the S&P 500 stocks. The S&P 500 index, which is a value-weighted sum of the most current stock prices, is subject to nonsynchronous trading problems. The index lags behind the true value of the underlying S&P 500 stocks when any of the constituent stocks have not recently traded since underlying stock values may change between trades. If the futures price is contemporaneously correlated with the underlying aggregate S&P 500 stock value, the index will lag behind the futures price. Nonsynchronous

* Associate Professor of Finance and Business Economics, University of Southern California.
Numerous individuals were very helpful in providing me the data necessary to undertake this study. They include J. Kimball Dietrich, of the U.S. Department of the Treasury (visiting), Kenneth Lehn, Chief Economist of the U.S. Securities and Exchange Commission, Bill MacDonald of the SEC, Ken Cone of the Chicago Mercantile Exchange, Jeremy Evnine of Wells Fargo Investment Advisors, and others unknown to me. Their assistance is greatly appreciated. The final version of the paper benefited from discussions in seminars at the SEC, the CFTC, the Fuqua School at Duke, and the Graduate School of Business at NYU. Other editorial suggestions were also provided by the editor, the referee, and Eduardo Schwartz. All remaining errors and all opinions are mine only. Partial financial support for this research was provided by the Mid America Institute for Public Policy Research, for which I am grateful.

trading also causes the index to be more autocorrelated and less volatile than the underlying aggregate stock value. If the nonsynchronous trading problem is ignored, spurious conclusions about volatility, market efficiency, and the relation between the futures and cash markets can be obtained and arbitrage opportunities can be falsely identified.

The nonsynchronous trading problem is greatest when prices are analyzed over short time intervals, such as are examined in studies of intraday data, and when trading is thin. Both conditions apply to detailed studies of market performance during the October 1987 crash. The crash probably will be best understood by studying intraday data, and, due to institutional factors and/or investor behavior, trading became thin at precisely those times when timely information would have been most useful. These considerations strongly suggest that better measures of the underlying values of the S&P 500 stocks be constructed and used. This paper derives such measures, compares them to the index, and uses them to analyze the relation between the two markets.

The new index measures are obtained directly from transactions data which identify the time and price of every trade. Briefly, each transaction price change is decomposed into a systematic component and a nonsystematic component. The systematic component is further decomposed into a sum of systematic price changes, one for each time interval spanned by the transaction price change. The systematic price changes, which are common to all transaction price changes spanning a given interval, are estimated as parameters using weighted sum of squares. The results are used to estimate what prices would have been observed had all securities traded. These are then used to compute an adjusted index. In effect, the estimation method obtains information from those stocks that did trade to determine the implied value of the underlying index.

The results suggest that part, but by no means all, of the large futures basis observed during the week of October 19–23 can be explained by nonsynchronous trading. The remainder may be due to disintegration (decoupling) of these two markets in which risk in the S&P 500 stocks is traded. The disintegration may have resulted when traders could not execute sell orders in the cash markets, when effective means of conducting arbitrage became unavailable (because of either order and confirmation channel congestion or regulation), and when the Chicago Mercantile Exchange halted trading in the futures contract.

Several studies have considered the effects of nonsynchronous trading on portfolio returns. These include Fisher [6], Dimson [4], Scholes and Williams [11], Cohen, Maier, Schwartz, and Whitcomb [2], Cohen, Hawawini, Maier, Schwartz, and Whitcomb [3], and, most recently, Atchison, Butler, and Simonds [1] and Shanken [12]. None of these studies, however, considers estimation solutions to nonsynchronous trading problems which can be executed using transactions data. Cohen et al. [2, 3] suggest that nonsynchronous trading may be only one of several market frictions which cause positive serial autocorrelation. Perry [10] and Atchison et al. [1] provide empirical and theoretical studies which imply that nonsynchronous trading, by itself, cannot account for the observed autocorrelation in daily index returns. The empirical analyses in this paper show that portfolio returns are autocorrelated even after the effects of nonsynchronous trading are explicitly removed.

Several studies consider the "causal" relation between stock index and index futures markets. Finnerty and Park [5], Kawaller, Koch, and Koch [7, 8], and Stoll and Whaley [13] all show that "causality" runs from the futures to the cash markets. However, as the authors are aware, the result may be caused by the lag that nonsynchronous trading induces into indices. This study shows that, even after the effect of nonsynchronous trading is removed, the futures price strongly leads the cash index.

MacKinley and Ramaswamy [9] provide a detailed theoretical and empirical study of the S&P 500 index-futures arbitrage conditions. They find that fifteen-minute changes in the computed index are more autocorrelated than changes in the futures and attribute this to nonsynchronous trading in the index stocks. Although they discuss how nonsynchronous trading can lead to the perception of arbitrage opportunities, they do not provide a method for eliminating the effect. This paper provides a method that can be implemented in real time.

The remainder of this paper consists of four sections. In Section I, the new method for computing indices from nonsynchronous data is motivated, derived, and presented. Section II describes the sample and experimental design of the empirical investigation into the S&P 500 index cash/futures relation surrounding October 19. Section III presents and describes the results. A summary with interpretations is provided in the last section, along with a discussion of the policy implications.

I. Methodology

Before presenting the weighted least-squares method for determining the aggregate market value of a stock portfolio from nonsynchronous trade prices, it is useful to introduce concepts and notation by discussing the standard definition of portfolio valuation. Portfolio value at time t, S_t, is normally computed as the sum over all N portfolio stocks of the number of shares held in each stock i, q_i, times the most recently observed price as of time t, P_{it}:

$$S_t = \sum_{i=1}^{N} q_i P_{it}.^{1} \tag{1}$$

For now, assume that prices, if observed, are observed at discrete intervals.

Let S_t^*, the aggregate value of the portfolio, be defined by

$$S_t^* = \sum_{i=1}^{N} q_i V_{it}, \tag{2}$$

where V_{it} is the value of a share of firm i at time t. Assume that this value equals the observed price when the price is observed. In general, share value differs from the last observed price if the latter is old. S_t^* is related to S_t as follows:

$$S_t^* - S_t = \sum_{i=1}^{N} q_i (V_{it} - P_{it}) = \sum_{i=1}^{N} q_i \Delta_{k_{it}} V_{it} \equiv A_t, \tag{3}$$

[1] The S&P 500 index is computed from (1) with $q_i = Q_i/DIV$, where Q_i is the total number of shares outstanding for stock i and DIV is a divisor published by S&P. The index is $1/DIV$ of the last-observed-price aggregate value of all S&P 500 stocks. It is a value-weighted index because the relative change in the index, $(S_t - S_{t-1})/S_{t-1}$, can be expressed as the value-weighted sum of the constituent stock returns.

where Δ_k is the k-period difference operator and k_{it} is the number of periods since the last price was observed for security i. ($k_{it} = 0$ if the price is observed at t.) If prices for all stocks are observed at time t, S_t is equal to S_t^*. Otherwise, S_t lags S_t^*. This expression shows that S_t^* can be estimated if A_t, the nonsynchronous trading adjustment, can be estimated.

The method presented here uses a simple one-factor representation of the value-generating process to estimate the nonsynchronous trading adjustment. Assume that values are generated by the following process:

$$\Delta \log(V_{it}) = f_t + e_{it}, \tag{4}$$

where f_t is the common factor and e_{it} is assumed to be zero-mean firm-specific variation. If a set of factor estimates $\{\hat{f}_t\}$ is available, unobserved multiperiod changes in value can be estimated by

$$\Delta_{k_{it}} \hat{V}_{it} = P_{it} \exp(\textstyle\sum_{i=1}^{k_{it}} \hat{f}_{t-i+1}), \tag{5}$$

where P_{it}, the most recently observed price as of time t, was observed k_{it} periods ago. Equations (3) and (5) reduce the problem of estimating S_t^* to one of estimating the underlying factor changes from nonsynchronous data.

The factor estimation method used in this study uses weighted least squares to allocate variance to the various intervals spanned by a price change. Motivation and interpretation of the method are best established by first considering a simple one-period problem.

The one-period percentage change in observed portfolio value, $\%\Delta S_t = (S_t - S_{t-1})/S_{t-1}$, can be computed by solving the following minimization problem:

$$\min_{f_t} \textstyle\sum_{i=1}^{N} w_i(\%\Delta P_{it} - f_t)^2, \tag{6}$$

where $w_i = q_i P_{it-1}/S_{t-1}$ is the value weight of stock i in the portfolio. The minimizing value of f_t, $\hat{f}_t = \sum w_i \%\Delta P_{it}/\sum w_i = \sum q_i \Delta P_{it}/S_{t-1}$ is identically equal to $\%\Delta S_t$,

This minimization problem has a weighted least-squares regression interpretation which is nearly identical to (4):

$$\%\Delta P_{it} = f_t + e_{it}, \qquad i = 1, \cdots, N, \tag{7}$$

where e_{it} is a firm-specific residual with variance proportional to $1/w_i$. This specification does not realistically represent the value-generating process (which may include multiple factors, non-unit factor loadings, and different firm-specific variance components); nor is it meant to. It is a consequence only of the definition of portfolio value. It does, however, suggest that factor estimates can be interpreted as percentage changes in portfolio value if value weights are used in weighted least-squares estimation.[2]

[2] More realistic factor models can be used in a nonsynchronous trading analysis within the framework presented in this article. In particular, multiple factors, non-unit loadings, realistic firm-specific variances, and information about bid/ask spreads can be incorporated. After examining the results obtained from the simple model presented here, it seems unlikely that a more complex model would yield results leading to different conclusions. Moreover, factor estimates obtained from more

Estimation of the factors from nonsynchronous data is accomplished using a multiperiod generalization of this simple regression model:

$$\Delta_{k_{it}} \log(P_{it}) = \sum_{j=1}^{K_{it}} f_{t-j+1} + e_{it}, \tag{8}$$

for all observed P_{it} in a cross-sectional sample of $i = 1, \cdots, N$ and a time-series sample of $t = t_0, \cdots, T$, with the variance of the firm-specific residual, e_{it}, proportional to $1/w_i$. If there are large numbers of independent observations at each time t, and if the price changes are small so that $\Delta\log(P)$ is approximately equal to $\%\Delta P$, $\hat{f}_t$ estimates $\%\Delta S_t^*$. (The latter condition is generally met in transactions data, even during the crash.)

The remainder of this section provides a brief discussion of several technical issues that arise when equation (8) is estimated. A complete presentation can be found in the Appendix.

Until now, all transactions, if observed, were assumed to occur at discrete intervals. In fact, they occur and are observed in continuous time. Failure to account for this characteristic of the data wastes information that can be used to obtain a better solution of the nonsynchronous trading problem. The Appendix describes and motivates a simple modification of equation (8) which allows extraction of this information from transaction prices that fall in the middle of an observational interval.

If there are a large number of intervals to be analyzed ($T - t_0$ is large), it is not practical to analyze them all at once. This would require computing and inverting a $T - t_0$ dimensional matrix of sums of squares and cross-products. Instead, a rolling regression spanning K intervals is used. The implementation is described fully in the Appendix.

Two types of nonsynchronous trading adjustments can be obtained from the rolling regressions: a current information adjustment and a perfect foresight adjustment. The current information adjustment for time t is computed using only the factor estimates of the regression whose leading edge is at t. This adjustment can be computed in real time using only those prices available at or before time t. The perfect foresight adjustment is computed using the factor estimates obtained from the middle interval of each rolling regression and therefore depends on future as well as past price changes. Although the perfect foresight measure may be the best measure of index value for testing certain hypotheses about investor behavior, other hypothesis tests requiring a measure which reflects only information available to investors as of time t are best examined using the current information adjustment. The empirical section describes the relation between the two measures and their differential implications for the future/cash basis.

Useful measures of uncertainty about index price changes and of uncertainty in index levels due to the nonsynchronous trading problem can be obtained from the inverse sum of squares and cross-products matrix of the regression. These measures, which quantify index precision and the loss in information due to nonsynchronous trading, are more fully described in the Appendix.

---

complex models cannot be interpreted as estimates of the change in portfolio value. (This property, though, is not necessary to the nonsynchronous trading analysis.)

II. Sample

The stock sample consists of all primary market trades of each S&P 500 stock from the open of trading on Monday, October 12, 1987 to the close of trading on Friday, October 23. The data, which come from SIAC via the Securities and Exchange Commission, include the date, time, price, and shares traded for each transaction on each exchange in the United States. The transaction time is the exact second when a trade was executed in the DOT system or when a record of it was read by an exchange card reader. Only primary market transactions are used since the S&P 500 index is computed only from primary prices. The primary markets for all S&P 500 stocks opened at 9:30 A.M. EST and closed at 4:00 P.M. except on the last sample day, when they all closed at 2:00 P.M. Some of the stocks traded later (except on Friday, October 23) on the Pacific Exchange. The NYSE is the primary market for 462 of the S&P 500 stocks, AMEX for eight, and the NASDAQ National Market System for the remaining 30. Shares outstanding (used to compute market weights) for each of the S&P 500 stocks are obtained from the October 1987 Standard and Poor's *"500" Information Bulletin*, and, where appropriate, the shares outstanding and/or prices are adjusted for all stock splits, stock dividends, new issues, and announced repurchases.[3]

The futures sample consists of all prices recorded by the CME market recorders in the December 1987 S&P 500 index for the corresponding sample period. Trade opened at 9:30 A.M. EST and closed at 4:15 P.M. EST except on the last day, when it closed at 2:00 P.M. EST. Each record includes the exact second that the price was entered by the recorder. On Tuesday, October 20, the exchange closed for about one hour near midday. Futures traders at the CME received reports of the S&P 500 index, computed from SIAC data, with no more than a thirty-second lag (even if the Composite Tape is running late).

The sample period is divided into 765 discrete intervals consisting of seventy-eight five-minute intervals for each full day (6.5 hours × 12 intervals/hour), fifty-four five-minute intervals for the last Friday, and nine overnight intervals. In the Appendix it is shown that only the first and last prices (if available) in each interval for a given stock are informative. Accordingly, the analysis is conducted using only these prices.

The first and last prices within an intraday interval are easy to identify, but two additional comments about the last and first daily intervals are necessary. First, although the exchanges nominally close at 4:00, many trades are recorded as late as 4:10 because trading in a crowd did not stop or because there are lags in the floor recording process. If the last trade of the day was recorded after 4:00, it is used as the last trade in the last interval and, for the purposes of the statistical analyses, its time is taken as 4:00. Second, the opening trade is often

[3] My computations of the S&P 500 index are consistently 0.3 points less than those reported in the November issue of Standard and Poor's *"500" Information Bulletin*. The difference almost certainly is due to inaccuracies in the number of shares outstanding. I obtained information about share changes from *Standard and Poor's Corporation Records, Current News Edition*, from the *Wall Street Journal*, and by comparing the numbers of shares outstanding reported in the October and November issues of the *"500" Information Bulletin*. Since the differences are small and consistent, they do not materially affect the analyses.

not recorded until some time after the nominal start of trade at 9:30. If a trade was recorded in the first five minutes, it is classified as the opening trade and its time is taken as 9:30. The first daily interval therefore covers the time between the first trade of the day (if recorded before 9:35) and 9:35. The overnight interval covers the time between 4:00 (or the last trade if recorded after 4:00) and the first trade of the day (if recorded before 9:35). To ensure that the futures data are compatible when overnight comparisons are made, the overnight change in the futures price is computed from the last price before 4:00 to the first opening price. (Both prices were recorded in the last and first minutes, respectively, on each trading day in the sample.)

Although some time series in the data set start at 9:30 on Monday, October 12, the first results concerning index levels reported in this paper start at interval 38 (Monday, October 12, 12:35–12:40). The delay occurs because index computation requires all prices. The last S&P 500 stock to open on that Monday first traded at 12:37. The regression analysis, which is designed to operate on incomplete data, uses all of the data starting from the Monday open. (Had closing prices for Friday, October 9 been available, the time-series results could have been extended back to the Monday open.)

III. Results

A. *Overview*

The CME December 1987 S&P 500 index futures contract price and the cash S&P 500 index are plotted by five-minute intervals in Figure 1. During the week

Figure 1. CME December 1987 S&P 500 index futures contract price (*solid line*) and the cash S&P 500 index (*broken line*), plotted by five-minute intervals.

of October 12–19, the two time series tracked each other very closely until Friday, when the futures started to trade at a discount to the cash. During the next week, the relation significantly deteriorated, with the basis ranging from −60 to +10 points. Two not necessarily exclusive explanations can account for the breakdown in the basis: nonsynchronous trading and disintegration.

If prices are nonsynchronous, the index will lag the underlying value of the S&P 500 stocks. (The lag is especially obvious on Figure 1 near the open of trading on October 19 and 20.) If that value is more currently reflected in futures price, and if it changes quickly, the basis will widen in the direction of change. This was generally observed between late Friday, October 16, and midmorning on Tuesday, October 20. Figure 2 presents time plots of two measures of trade frequency: the value-weighted mean number of intervals (and fractions of intervals) since the last trade was observed for stocks on the S&P 500 list, and the fraction of the total market value of the S&P 500 stocks which did not trade in a given interval. The time periods in which trading is least frequent correspond to those with the greatest absolute basis.

Under the disintegration hypothesis, the basis breaks down when mechanisms that integrate these two markets for fundamentally the same risk fail. This could have been the case during much of the week of October 19–23, when traders could not always execute sell orders in the cash markets, when effective means of conducting index arbitrage were unavailable (either because of order and

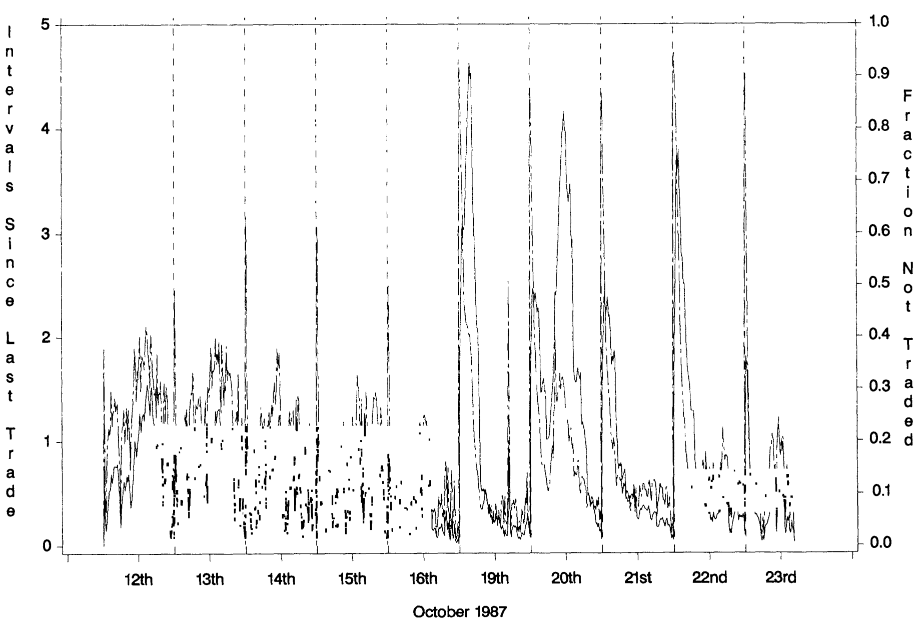


Figure 2. Value-weighted mean number of intervals, including fractional intervals, since the last trade was observed for stocks on the S&P 500 list (*solid line*) and the fraction of the total market value of the S&P 500 stocks which did not trade in a given interval (*broken line*). Each interval is five minutes long, except for the first interval of each day, which corresponds to the period between the preceding 4:00 close and the opening transaction, if observed before 9:35.

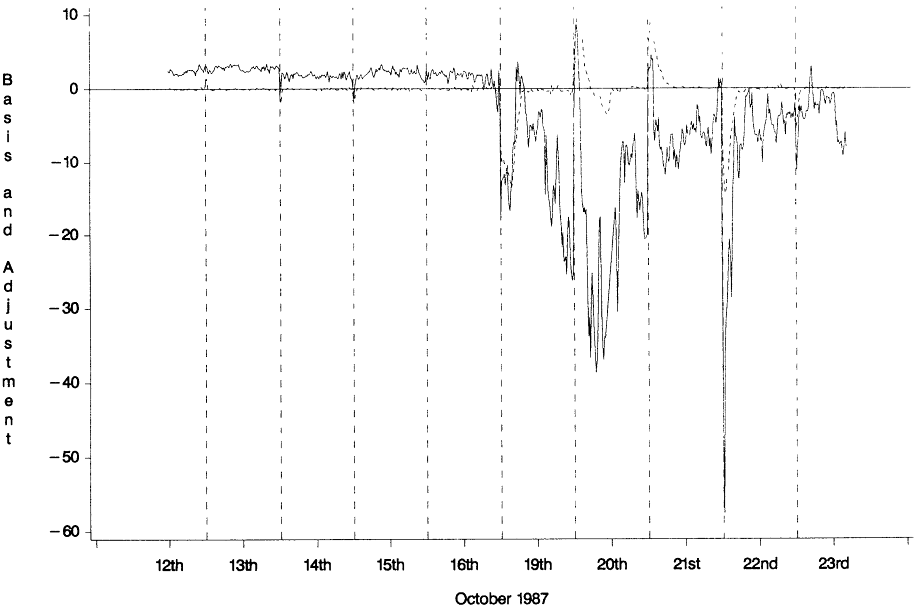


Figure 3. CME December 1987 S&P 500 index futures basis (*solid line*) and the perfect foresight nonsynchronous index adjustment (*broken line*), plotted by five-minute intervals. The current information adjustment, if plotted, would lie almost exactly on the perfect foresight adjustment. The basis is futures price minus the cash index.

conformation channel congestion or regulation), and when the CME halted trading in the futures contract.

The methods developed in this paper provide the tools for measuring how much of the abnormal basis was due to nonsynchronous trading. Time plots of the basis (futures minus cash) and of the perfect foresight nonsynchronous index adjustment are presented in Figure 3. (The current information adjustment, if plotted, would lie almost exactly on the perfect foresight adjustment.[4]) If nonsynchronous trading explained all of the variation in the basis, the basis and the adjustment would track each other very closely at a distance of about 2.5 points. (The actual distance depends on expected dividend yields and carrying costs.) The plots show that the adjustment explains some of the variation in the basis, especially near market openings. Much of the remaining variation, however, remains unexplained. It may be due to market disintegration.

Figure 4 plots the estimated standard errors of the perfect foresight index adjustment and of the estimated change in the S&P 500 index implied by the perfect foresight estimated factor. As noted in the Appendix, the standard errors obtained from the regression analysis are not the actual standard errors of estimators (since the variance weights used in the regression do not correspond

[4] In all of the results reported, there is very little difference between the current information adjustment and the perfect foresight adjustment. This implies that the cross-section is large enough that estimation error is not a significant source of time-series variation in the factor estimates. Current information alone appears to be sufficient to accurately estimate the factors in this sample.

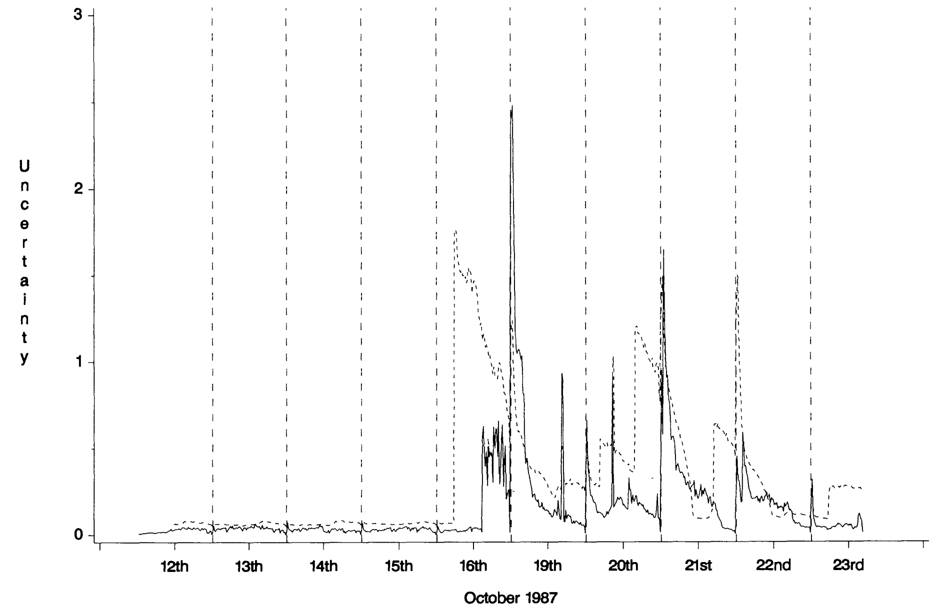


Figure 4. Estimated standard error of the perfect foresight index adjustment (*solid line*) and of the estimated change in the S&P 500 index, implied by the perfect foresight estimated factor (*broken line*). These standard errors obtained from the regression analysis are not the actual standard errors of estimators since the variance weights used in the regression do not correspond to the actual firm-specific component variances.

to the actual firm-specific component variances). They do, however, show that uncertainty due to nonsynchronous trading occurs when it is expected, given the time-since-last-trade data plotted in Figure 2.

B. Correlations, Lead-Lag Relations, and Autocorrelations

Quantitative measures of the tracking performance of the index and the adjusted indices are presented in Table I. The results confirm that the adjusted indices track the futures better than does the S&P 500 index. When prices are most volatile, the perfect foresight-adjusted index is generally closest to the futures price, the current information-adjusted index is second-closest, and the index is farthest away. Otherwise, all three measures of the index are nearly coincident. (If prices are nonsynchronous, the index will accurately represent value if value does not change very quickly.) Although the mean differences are not large, paired t-tests show that the mean basis distances for the adjusted indices are generally significantly smaller than for the index. The differences between the perfect foresight and the current information indices are also often significant.

Further evidence suggesting that the adjusted indices more closely approximate the aggregate underlying value of the S&P 500 stocks than does the S&P 500 index can be found by examining correlations. Since the futures contract and the S&P 500 index both price the same risks, changes in their values should be

Table I

Distance Measures of the Basis (Futures-Index) between the CME December 1987 S&P 500 Index Futures and the S&P 500 Index, and the Two Adjusted Indices[a]

		Mean Basis			Mean Absolute Basis			Root Mean-Squared Basis		
Sample	N	Index	Current Adjusted	Perfect Adjusted	Index	Current Adjusted	Perfect Adjusted	Index	Current Adjusted	Perfect Adjusted
All	717	−3.60	−3.36	−3.31	5.88	5.62[b]	5.57[bc]	9.44	8.91[b]	8.83[bc]
Week 1	356	2.07	2.10	2.11	2.12	2.13	2.14[c]	2.22	2.22	2.23[c]
Week 2	361	−9.20	−8.75	−8.65	9.60	9.06[b]	8.96[bc]	13.12	12.37[b]	12.25[bc]
12th	40	2.36	2.36	2.37	2.36	2.36	2.37	2.39	2.38	2.39
13th	79	2.79	2.77	2.77	2.79	2.77	2.77	2.81	2.79	2.79
14th	79	1.70	1.75	1.76	1.72	1.75	1.76	1.77	1.78	1.79
15th	79	2.07	2.14	2.16	2.11	2.14	2.16	2.20	2.21	2.22
16th	79	1.56	1.60	1.60	1.73	1.75	1.76	1.86	1.86	1.87
19th	79	−10.55	−8.18	−7.83	10.89	9.30[b]	9.04[bc]	13.08	11.84[b]	11.79[bc]
20th	69	−16.99	−18.00	−18.10	17.88	18.12	18.14	20.61	20.70	20.67
21st	79	−5.57	−6.68	−6.74	6.15	6.76[b]	6.81[b]	6.68	7.40[b]	7.43[b]
22nd	79	−8.76	−7.30	−7.08	8.76	7.30[b]	7.08[bc]	13.65	10.89[b]	10.39[bc]
23rd	55	−3.30	−3.00	−2.96	3.51	3.16[b]	3.13[b]	4.44	3.99[b]	3.96[b]

[a] The current information-adjusted index is the index plus an adjustment for nonsynchronous trading that is computed from lagged and contemporaneous information only. The adjustment for the perfect foresight-adjusted index is obtained using leading data as well. The measures are computed from data observed at five-minute intervals starting at market open.

[b] Mean significantly different from the corresponding index basis mean using a paired t-test with a one-sided significance level of five percent.

[c] Mean significantly different from the current adjusted basis mean using a paired t-test with a two-sided significance level of five percent.

correlated. Nonsynchronous trading in the cash market, however, tends to break up that correlation. The adjusted indices therefore should be more highly correlated with changes in the futures price than is the index. Results reported in Panel A of Table II confirm this prediction. For the sample as a whole and every time subsample, both adjusted measures are more closely correlated with the futures than is the index.

Information about the time relation between the futures and cash markets can be obtained by examining cross-correlations. Modeling nonsynchronous trading is very important in this analysis because, even if the relation is completely contemporaneous, nonsynchronous trading will cause the futures price to lead the computed index. This argument suggests that the lagged changes in the futures should be more highly correlated with current changes in the index than with those of the adjusted indices. However, if in the true causal relation, the futures leads the cash (as might be expected given the lower transactions costs and greater liquidity in the futures markets), lagged futures should be less highly correlated with the current index than with the current adjusted indices. This is because the index contains noise induced by nonsynchronous trading. The results (Table II, Panel B and Figure 5) show that the futures leads all measures of the index. In the first week of the sample, when noise due to nonsynchronous trading is not great, the lead correlations (negative lags of the futures) are greatest for the S&P 500 index, as predicted. In the second week, the lead correlations are initially greatest for the adjusted indices, consistent with the noise prediction. There is little evidence of the cash leading the futures in any of the measures (Table II, Panel C).

An interesting implication of the cross-correlation between the current changes in the various indices and the leads and lags of changes in the futures is that lagged adjustments have no predictive power for futures prices. This can be seen most clearly in Figure 6, in which the sample cross-correlation functions are plotted for current values of the two adjustments with leads and lagged changes in the futures. These results show that the futures market "sees through" (is efficient with respect to) noise due to nonsynchronous trading.

One prediction of the nonsynchronous trading analysis is that the S&P 500 index should be positively autocorrelated. In five-minute data (Table III and Figure 7), it is very autocorrelated (0.697 for the whole sample) and, as expected, the adjusted indices are less correlated. The autocorrelation in the adjusted series, however, is still quite large (0.527 for perfect foresight adjustment) relative to the autocorrelation in the futures (0.143). These results suggest that many stock specialists attempted to maintain price continuity. They also show that serious problems would result if the nonsynchronous trading problems were modeled using time-series methods which assume that the underlying index value follows a random walk.

The nonsynchronous data analysis predicts that the index should catch up with its underlying value as trades are observed. This implies that the adjustments should lead the index. The results (Table IV and Figure 8) confirm this prediction. The current adjustments are correlated with the lagged and current values of the index change (because the adjustments are computed partly from lagged and current index changes) and with leads of the index change, as expected. Note

Table II

Correlations of Changes in the S&P 500 Index, the Adjusted Indices, and the Index-scaled Estimated Factors, with Changes in the CME December 1987 Futures Price[a]

Sample	N	S&P 500 Index	Index with Current Information Adjustment	Index with Perfect Foresight Adjustment	Estimated Factor Change, Current Information	Estimated Factor Change, Perfect Foresight
Panel A: Correlations with Contemporaneous Changes in the Futures Price						
All	716	0.196	0.347	0.444	0.371	0.431
Week 1	356	0.587	0.682	0.690	0.683	0.671
Week 2	360	0.172	0.332	0.438	0.359	0.425
12th	40	0.670	0.711	0.732	0.761	0.725
13th	79	0.667	0.719	0.739	0.751	0.743
14th	79	0.528	0.737	0.750	0.753	0.768
15th	79	0.633	0.734	0.769	0.771	0.772
16th	79	0.574	0.654	0.641	0.615	0.587
19th	79	0.182	0.393	0.429	0.428	0.405
20th	68	0.177	0.296	0.367	0.304	0.332
21st	79	0.249	0.576	0.641	0.607	0.656
22nd	79	0.039	0.195	0.444	0.246	0.448
23rd	55	0.355	0.639	0.630	0.660	0.583
Panel B: Correlations with First Lagged Changes in the Futures Price						
All	715	0.259	0.403	0.391	0.394	0.369
Week 1	355	0.531	0.421	0.434	0.433	0.426
Week 2	360	0.248	0.411	0.397	0.402	0.372
12th	39	0.726	0.544	0.550	0.583	0.540
13th	79	0.328	0.213	0.216	0.239	0.200
14th	79	0.300	0.298	0.329	0.307	0.322
15th	79	0.614	0.496	0.509	0.505	0.488
16th	79	0.610	0.477	0.484	0.478	0.482
19th	79	0.364	0.417	0.434	0.435	0.381
20th	68	0.323	0.489	0.502	0.491	0.456
21st	79	0.203	0.399	0.325	0.400	0.318
22nd	79	0.057	0.406	0.370	0.358	0.355
23rd	55	0.344	0.138	0.178	0.203	0.170
Panel C: Correlations with First Lead Changes in the Futures Price						
All	715	0.001	0.040	0.047	0.045	0.030
Week 1	355	0.005	0.055	0.033	0.040	−0.013
Week 2	360	0.000	0.039	0.049	0.046	0.033
12th	39	0.352	0.454	0.390	0.405	0.266
13th	79	−0.143	0.069	0.027	0.039	−0.017
14th	79	−0.189	−0.173	−0.201	−0.183	−0.248
15th	79	0.201	0.224	0.218	0.238	0.179
16th	79	−0.066	−0.007	−0.026	−0.025	−0.072
19th	79	0.033	−0.015	0.001	0.013	−0.036
20th	68	−0.062	0.069	0.064	0.061	0.046
21st	79	−0.170	−0.005	−0.004	−0.038	−0.033
22nd	79	−0.025	−0.043	−0.023	−0.029	−0.038
23rd	55	0.046	0.285	0.300	0.340	0.306

[a] The changes for each day are computed over five-minute intervals with the preceding overnight change included.

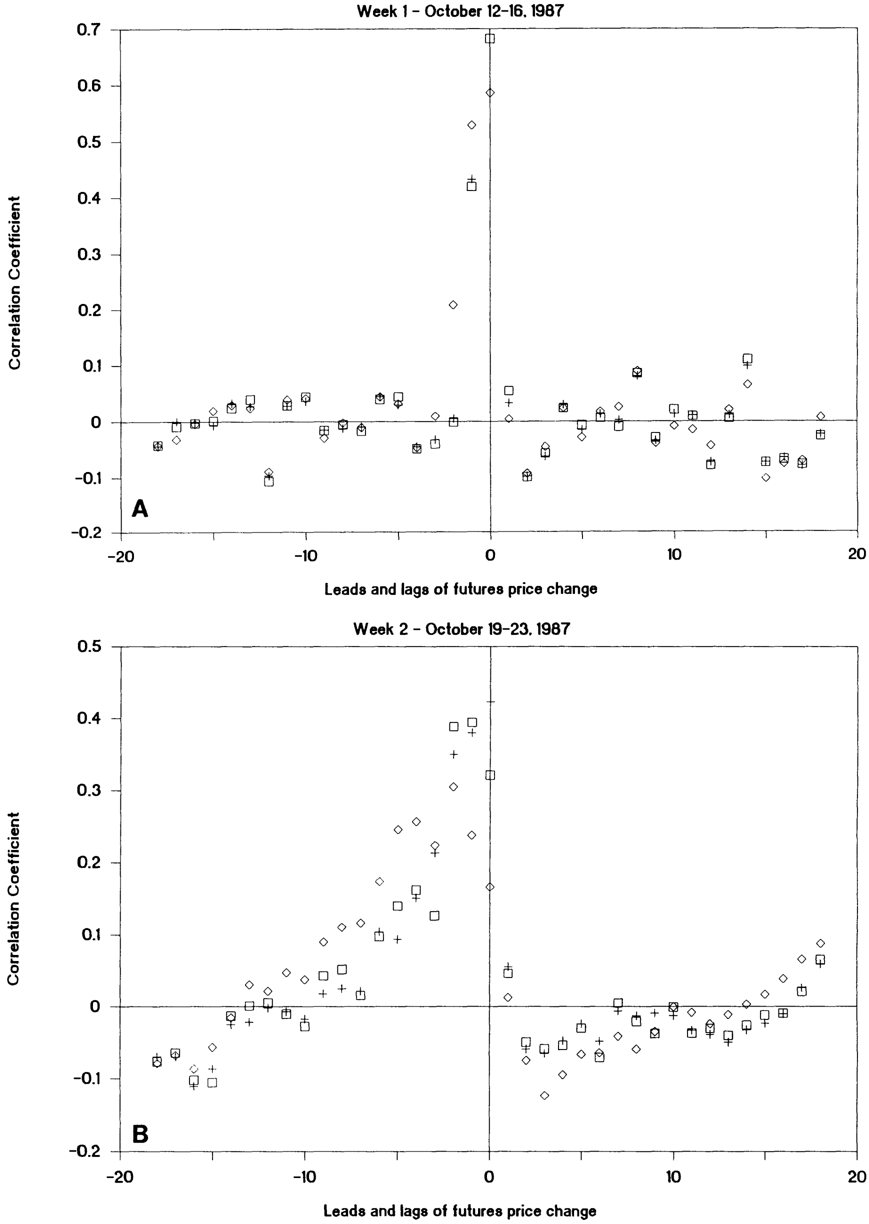


Figure 5. Cross-correlations, current five-minute changes in the S&P 500 index and in the two adjusted indices with leads and lags of changes in the CME December 1987 S&P 500 index futures. There are 356 five-minute and overnight intervals plotted. Positive correlations to the left of zero indicate that the futures leads the indices. The *diamonds* plot correlations of S&P 500 index changes with leads and lags of changes in the futures price. The *boxes* plot cross-correlations for changes in the current information foresight index, while the *pluses* plot cross-correlations for changes in the perfect foresight index.

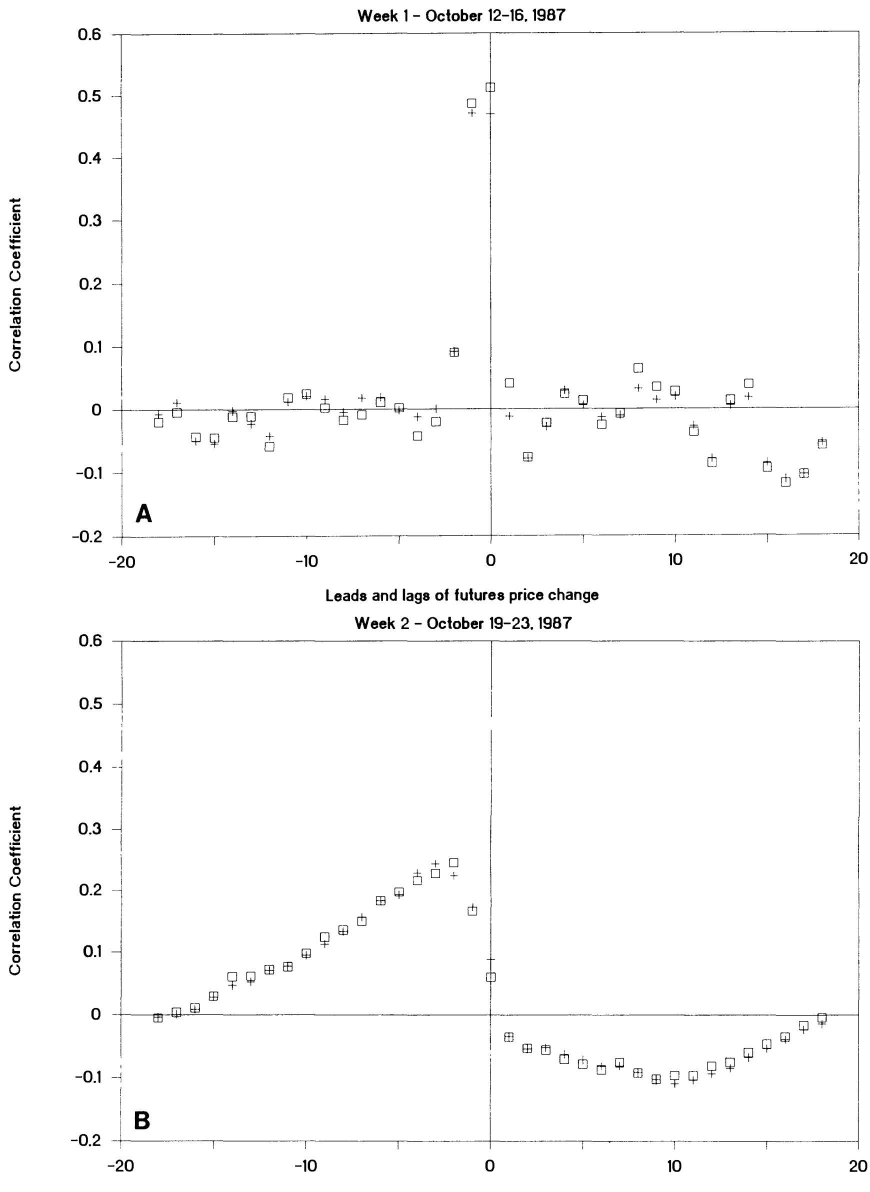


Figure 6. Cross-correlations, current values of the two nonsynchronous trading adjustments with leads and lags of five-minute changes in the CME December 1987 S&P 500 index futures price. There are 356 five-minute and overnight intervals plotted. Positive correlations to the left of zero indicate that the futures leads the adjustments. The *boxes* plot the perfect foresight adjustment correlations, while the *pluses* plot the current information adjustment correlations.

Table III

First Autocorrelation Coefficients for Changes in the CME December 1987 S&P 500 Index Futures Price, the S&P 500 Index, the Adjusted Indices, and the Index-Scaled Estimated Factors[a]

Sample	N	CME December 1987 Futures	S&P 500 Index	Index with Current Information Adjustment	Index with Perfect Foresight Adjustment	Estimated Factor Change, Current Information	Estimated Factor Change, Perfect Foresight
All	726	0.143	0.697	0.458	0.527	0.562	0.559
Week 1	356	0.059	0.372	0.190	0.205	0.234	0.151
Week 2	371[b]	0.145	0.755	0.496	0.574	0.613	0.614
12th	39	0.290	0.635	0.385	0.406	0.477	0.343
13th	79	−0.065	0.026	0.054	0.051	0.097	0.011
14th	79	−0.082	0.154	0.013	0.034	0.073	0.022
15th	79	0.338	0.398	0.132	0.190	0.207	0.170
16th	79	0.001	0.454	0.292	0.278	0.288	0.167
19th	79	−0.025	0.554	0.248	0.325	0.383	0.426
20th	79[b]	0.286	0.873	0.819	0.851	0.873	0.850
21st	79	−0.030	0.654	0.422	0.386	0.377	0.401
22nd	79	0.132	0.777	0.379	0.645	0.602	0.644
23rd	55	0.195	0.447	0.107	0.135	0.198	0.121

[a] The changes for each day are computed over five-minute intervals with the preceding overnight change included.

[b] The futures sample is eleven observations smaller due to the halt of trading near midday on Tuesday, October 20.

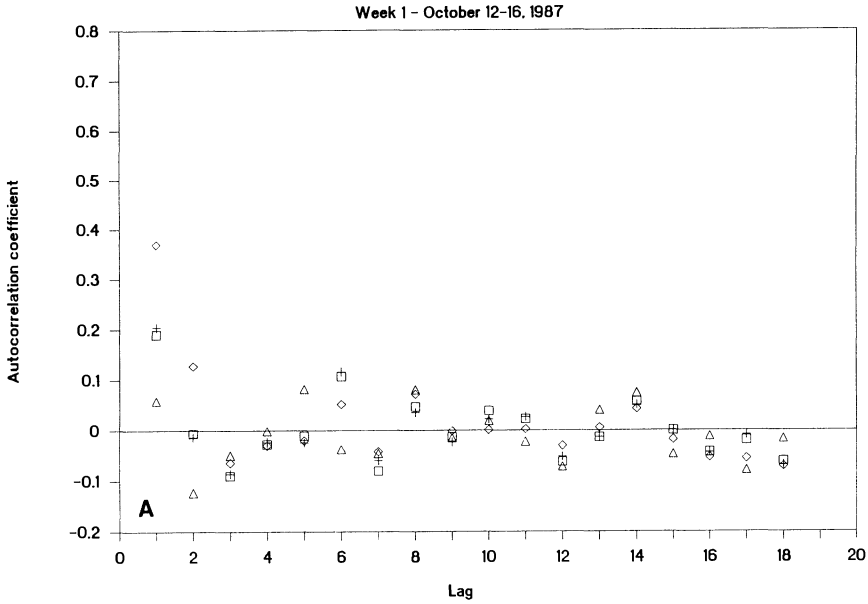
that this does not imply that the adjustments predict future changes in value. They predict past changes in value which have not yet been incorporated into the observed index.

Further evidence of the nonsynchronous trading-induced lags in the index can be found in Figure 9, in which the root mean-squared cross-basis function is plotted. The cross-basis is the difference between the index and a lagged or lead value of the futures. The minimum of this function indicates at what lag the relation between the futures and the index is closest. In the first week of the sample, this relation was closest for the contemporaneous basis. It is quite flat because prices did not change very much that week (Friday excepted). In the second week the minimum for the index is found at lag 5 of the futures and between lags 2 and 3 for the adjusted indices. This shows that the adjustments take out some of the lag observed in the index during the week, but not all.

IV. Summary, Interpretations, and Conclusions

This paper derives new estimators of the underlying value of a stock portfolio which abstract from nonsynchronous trading problems by using the complete transaction history of all stocks in the portfolio. The methods are applied to the S&P 500 index for a ten-day period surrounding the October 1987 stock crash.

Nonsynchronous trading can explain part of the large absolute futures-cash basis observed during the crash, but not all of it. Much of the unusually large



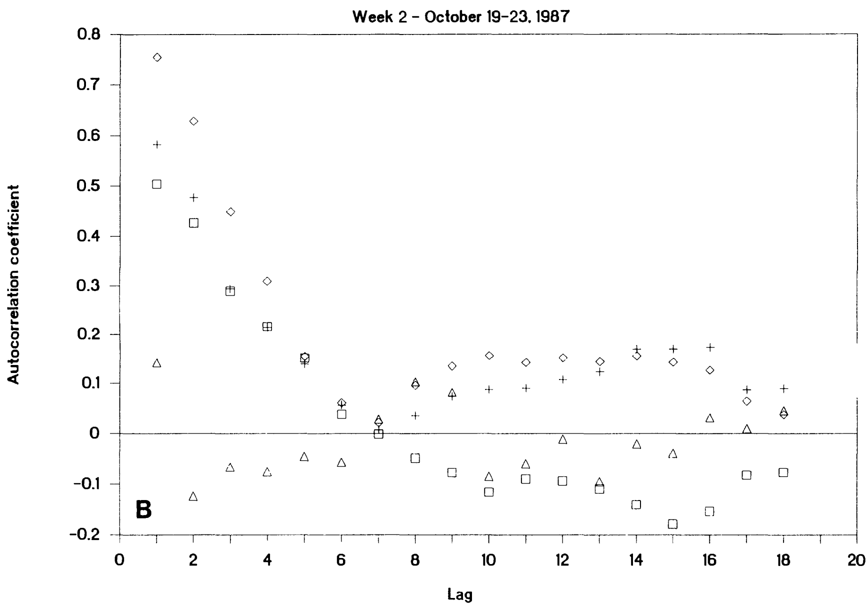


Figure 7. Autocorrelation coefficients for five-minute changes in the CME December 1987 S&P 500 index futures (*triangles*), in the S&P 500 index (*diamonds*), in the perfect foresight-adjusted index (*boxes*), and in the current information adjustment (*pluses*). There are 356 five-minute and overnight intervals plotted.

basis may be due to the disintegration of the two markets caused by capacity and/or regulatory disruptions in the trade processes.

The large index autocorrelations observed in the second week, even after adjustment for serial correlation, and the long lagged relation between the index

Table IV

Correlations of the Current Information and Perfect Foresight Adjustments to the S&P 500 Index with Changes in the Index[a]

		First Lagged Index		Contemporaneous Index		First Lead Index	
Sample	N[b]	Current	Perfect	Current	Perfect	Current	Perfect
All	727	0.473	0.449	0.516	0.485	0.474	0.457
Week 1	356	0.296	0.294	0.711	0.643	0.482	0.423
Week 2	371	0.506	0.481	0.540	0.510	0.501	0.487
12th	40	0.321	0.348	0.758	0.784	0.660	0.596
13th	79	0.364	0.353	0.699	0.693	0.414	0.392
14th	79	0.205	0.214	0.631	0.561	0.458	0.415
15th	79	0.358	0.365	0.737	0.638	0.517	0.485
16th	79	0.243	0.253	0.873	0.893	0.518	0.418
19th	79	0.252	0.252	0.286	0.276	0.230	0.234
20th	79	0.606	0.529	0.588	0.512	0.506	0.450
21st	79	0.636	0.618	0.749	0.751	0.733	0.748
22nd	79	0.471	0.466	0.536	0.517	0.530	0.543
23rd	55	0.312	0.304	0.436	0.421	0.377	0.372

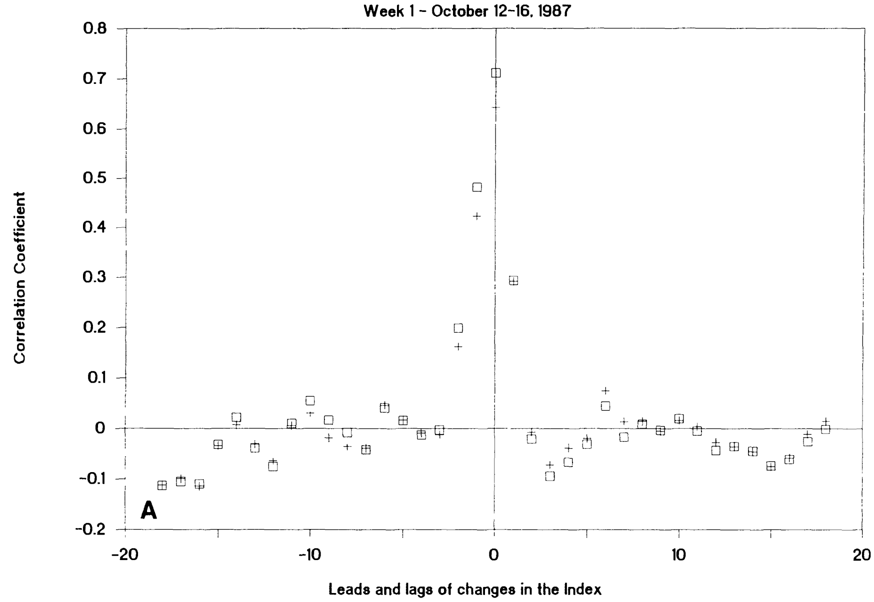
[a] The changes for each day are computed over five-minute intervals with the preceding overnight change included.

[b] Number of observations for the contemporaneous correlation. The lead and lag correlations have one less observation for the sample as a whole, for week 1 and for October 12.

and the futures during this week suggest that the futures market leads the cash market and that the cash market is not as efficient (as measured by serial correlations) over short intervals as is the futures market. Alternatively, these results may be interpreted as evidence that specialists were providing price continuity and that its provision took time.

These results may be more compelling than similar ones obtained in earlier studies since this study explicitly controls for the lagged cross-correlations and serial correlations which nonsynchronous trading introduces into the data. This study analyzes a very short sample period, but the results may be very strong because of the extraordinary variation in that sample.

A full analysis of why and how the crash occurred requires a much broader study than that presented here. However, the partial elimination of nonsynchronous trading as an explanation for the large basis reduces the set of possible causes and clarifies the problem. A reasonable interpretation of the results suggests that the crash might not have been as large as it was had more orderly trade mechanisms been maintained. In particular, the partial elimination of futures arbitrages due to exchange regulation, to congestion in the order and confirmation systems, and to other difficulties associated with executing sale orders in the cash market removed a potentially significant flow of buy orders from the futures market. Moreover, these same difficulties may have increased the number of sell orders coming into the futures market even though it traded at a significant discount. These two factors may have caused larger drops in the futures than might otherwise have been observed. This may have had a very significant effect



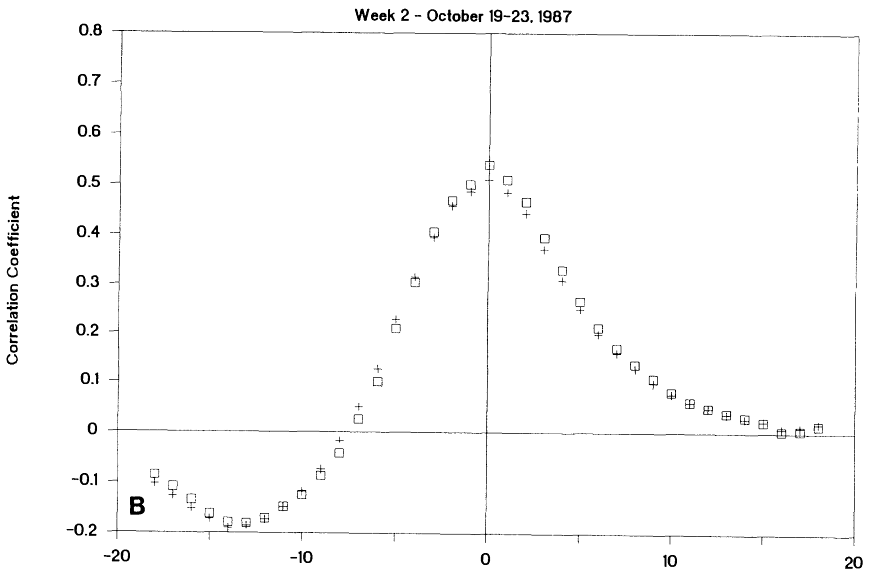


Figure 8. Cross-correlations, current values of the two nonsynchronous trading adjustments with leads and lags of five-minute changes in the S&P 500 index. There are 356 five-minute and overnight intervals plotted. Positive correlations to the right of zero indicate that the adjustments lead the index. The *boxes* plot the perfect foresight adjustment correlations, while the *pluses* plot the current information adjustment correlations.

on the overall performance of the two otherwise integrated markets since the evidence strongly suggests that the cash market follows the futures market.

This interpretation of the results suggests that future problems can be at least partially eliminated if capacity limits to the flow of information are raised. In

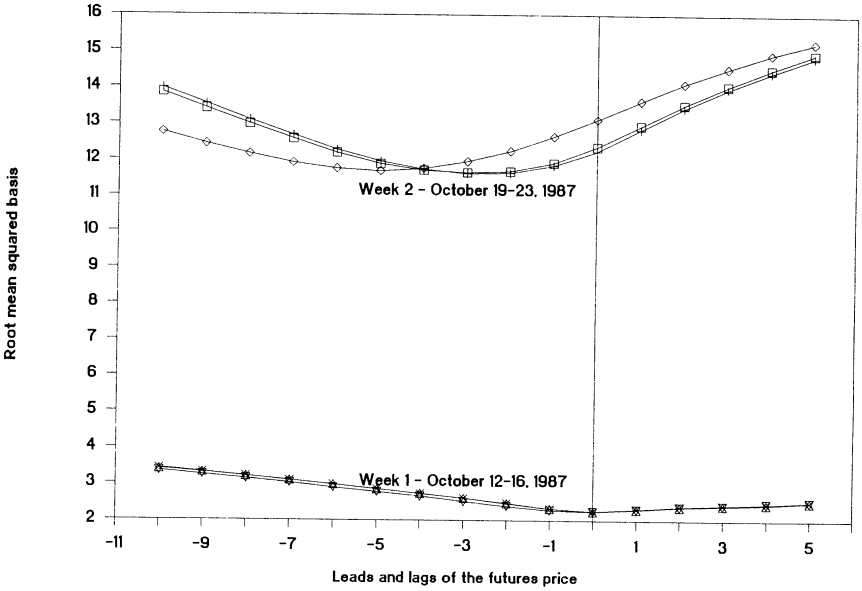


Figure 9. Root mean-squared cross-basis. The cross-basis is the difference between the index and a lead or lag of the futures price. The *diamonds* plot the cross-basis computed using the S&P 500 index, the *boxes* plot the cross-basis computed using the perfect foresight-adjusted index, while the *pluses* plot the cross-basis computed using the current information-adjusted index.

particular, orders, confirmations, and transaction reports on the ticker tape should all arrive at their destinations instantly.

Also, the requirement that specialists maintain continuous markets should be relaxed somewhat in the direction of broad market moves. In practice, this could be done by monitoring the difference between the percentage price change in a specialist's stock and the percentage change in some market index over the same time interval. This relaxation of the requirement to maintain continuous markets should be offset by requiring specialists to quote tighter spreads.

Appendix: Methodological Details

A. *Intra-Interval Prices*

The discussion in Section I assumes that transactions, if observed, occur at discrete intervals. In fact, they occur and are observed in continuous time. This can cause problems when estimating (8). Suppose that the regression is estimated using only the last transaction price, if available, from each of the $T - t_0$ time intervals. This simple and common practice destroys useful information concerning when within the interval the price is observed. For example, let measurement intervals be five minutes long and suppose that only two prices are observed in intervals $t - 2$ to t: one at 0.5 minutes into interval $t - 2$ and the other at one minute into interval t. It would be inappropriate to allow the implied price change to equally influence the estimation of f_{t-1} and f_t, as would happen using this simple practice. Doing so would allocate too much variation to interval t, which is hardly spanned by the price change, and none to interval $t-2$, which is mostly spanned by the price change. This problem can be partly solved by allocating variance in proportion to the span of the observed price change over that interval. For the above example, (8) would be modified to read: $(P_t - P_{t-2})/P_{t-2} = 0.9f_{t-2} + f_{t-1} + 0.2f_t + e_t$. This allocation method can be justified if we assume that prices within an interval are generated by Gaussian diffusion with constant variance.[5]

Although this allocation method would permit all transaction prices to be used in the estimation, it is best not to use them all since doing so unnecessarily increases estimator variance. All information about the total factor change within a given interval is contained only in the first and last transactions in that interval. The use of intermediate transactions increases noise because the intra-interval factor variance decomposition is done with error; it is assumed to be proportional to the total factor change over the interval.[6] Therefore, the model is estimated using only first and last interval prices, if available.[7]

[5] Suppose that f^b, f^a, and f are random variables such that $f^b + f^a = f$ with f^b and f^a uncorrelated. Interpret f as the change in the factor over the entire interval, f^b as the change in the factor before a price is observed, and f^a as the change in the factor after the price is observed. We want to know the conditional mean of f^b given f. It is $\mathrm{E}(f^b|f) = \mathrm{cov}(f^b, f)/\mathrm{var}(f)\ f$, which can be reduced to $\mathrm{var}(f^b)/\mathrm{var}(f)\ f$ since f^b and f^a are uncorrelated. If the factor evolves at a uniform rate, the ratio of variances is equal to the fraction of time in the interval past when the price is observed.

[6] Using the intermediate price increases estimation error because the intra-interval variance is poorly modeled. This can be illustrated in the following example. Suppose that price change d_1 spans the subinterval from 0 to x and price change d_2 spans the subinterval from x to 1, $0 < x < 1$. Let these price changes be generated from the following model:

$$d_1 = f_1 + e_1,$$

$$d_2 = f_2 + e_2,$$

where f_1, f_2, e_1, and e_2 are uncorrelated and e_1 and e_2 have variances respectively equal to x and $(1-x)$ and f_1 and f_2 have variances respectively equal to sx and $s(1-x)$, where s is some positive constant. Consider the problem of estimating $f = f_1 + f_2$ given d_1 and d_2. The variance decomposition procedure implies the following regression model:

$$d_1 = xf + n_1,$$

$$d_2 = (1-x)f + n_2.$$

The OLS estimator of f is $\hat{f} = b_1 d_1 + b_2 d_2$, where $b_1 = x/[x^2 + (1-x)^2]$ and $b_2 = (1-x)/[x^2 + (1-x)^2]$. The mean-squared error in this estimator has two components. The first component is due to the fact that f_1 and f_2 are not equal to xf and $(1-x)f$, respectively. This component is equal to $(b_1^2 - 1)sx + (b_2^2 - 1)s(1-x) + b_1^2 x + b_2^2(1-x)$. The second component is due to the noise caused by e_1 and e_2. This component is equal to $[x^3 + (1-x)^3]/[x^2 + (1-x)^2]^2$. For all x, this second component is greater than or equal (at $x = 0.5$) to one, which is the total estimator mean-squared error of $\hat{f} = d_1 + d_2$, the best estimator.

[7] An additional example provides further intuition for why the first transaction in each interval should be included in the analysis. Suppose that prices are observed at the beginning of interval $t-1$, 0.5 minutes into interval t, and at the end of interval t. The intermediate transaction, which would be ignored if we only examined the last transaction in each interval, clearly contains useful information about the decomposition of total variance between the two intervals.

B. Uncertainty Measures

Useful measures of uncertainty can be obtained from the estimated variance-covariance matrix of the regression model estimates. The estimated standard error of $\hat{f}_t$, however, must be interpreted with caution since it reflects the value weights used in the estimation and not the actual variance of the firm-specific variance components. It is, however, positively related to the extent of nonsynchronous trading and is therefore useful as an ordinal measure of factor uncertainty due to nonsynchronous trading.[8] An ordinal measure of uncertainty in the index due to nonsynchronous trading can also be obtained from the estimator variance-covariance matrix since the nonsynchronous trading adjustment, A_t, computed using (3) and (5), is approximately a linear sum of the factor change estimates. The approximation error is very small since the second and higher order terms in the expansion of the exponential function in (5) are very small. The estimated standard error of this sum is a measure of the uncertainty in the index due to nonsynchronous trading.

C. Rolling Regressions

Since the number of intervals to be analyzed is large, it is not practical to analyze them all at once. Instead, a rolling regression spanning K intervals is used. In order to preserve as much information as possible, the rolling regression is started at interval 1, spanning only that interval. In the next step, intervals 1 and 2 are analyzed and, in the next, intervals 1 to 3. The analysis continues in this manner until the regression spans intervals 1 to K, after which the next regression spans intervals 2 to $K + 1$. The regression then steps through the data until intervals $T - t_0 - K$ to $T - t_0$ are analyzed, at which point the analysis is finished.[9]

D. Estimator Precision

Simulated data were created and analyzed using this new econometric method. The results (not reported) confirm that the machine algorithms are accurate and that the method is able to recover underlying factors from nonsynchronous data, even when the data are far less synchronous than actual stock market prices. There is, however, an upper bound on estimator precision as the cross-sectional sample increases for a given interval length. The bound is due to unmodeled intra-interval variation associated with observations which arrive in the middle of an interval. The proportional allocation rule for this variation proposed above

[8] If cardinal measures are desired, they can be computed from estimates of the variance of firm-specific variance components since the weighted least-squares estimator is a linear function of the data. Note also that, if there is no nonsynchronous trading, the portfolio values would be known with certainty but not the underlying factor values. Uncertainty about the factor values depends on the number of stocks in the portfolio and on their weighting within the portfolio.

[9] The rolling regressions use a span of sixty-one intervals. A comparison of results from smaller regressions spanning a maximum of fifteen and thirty-one intervals suggests that little would be gained from using a longer span. The factor estimate for interval t used for computing perfect foresight adjustments is taken from the 61st rolling regression if $t \leq 31$, from the t + 30th rolling regression if $31 \leq t \leq 735$, and from the 735th regression if $t \geq 735$.

raises, but cannot fully remove, this bound. If the cross-sectional sample is large, overall precision (per unit of time) can be increased by decreasing the analysis interval length. The simulation results and the results reported above showing near equivalence of the two adjustment methods suggest that in this study it is unlikely that estimator noise is a large source of variation relative to variation due to the underlying factor.

REFERENCES

1. Michael D. Atchison, Kirt C. Butler, and Richard B. Simonds. "Nonsynchronous Security Trading and Market Index Autocorrelation." *Journal of Finance* 42 (March 1987), 111–18.
2. Kalman J. Cohen, Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb. "On the Existence of Serial Correlation in an Efficient Securities Market." *TIMS Studies in the Management Sciences* 11 (1979), 151–68.
3. Kalman J. Cohen, Gabriel Hawawini, Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb. "Friction in the Trading Process and the Estimation of Systematic Risk." *Journal of Financial Economics* 12 (August 1983), 263–78.
4. E. Dimson. "Risk Measurement When Shares Are Subject to Infrequent Trading." *Journal of Financial Economics* 7 (June 1979), 197–226.
5. Joseph E. Finnerty and Hun Y. Park. "Does the Tail Wag the Dog?: Stock Index Futures." *Financial Analysts Journal* 43 (March 1987), 57–58.
6. Lawrence Fisher. "Some New Stock Market Indexes." *Journal of Business* 39 (January 1966), 191–225.
7. Ira G. Kawaller, Paul D. Koch, and Timothy W. Koch. "The Extent of Feedback Between S&P 500 Futures Prices and the S&P 500 Index." Working paper, Chicago Mercantile Exchange, 1987.
8. ———. "The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index." *Journal of Finance* 42 (December 1987), 1309–29.
9. A. Craig MacKinlay and Krishna Ramaswamy. "Index-Futures Arbitrage and the Behavior of Stock Index Futures Prices." *Review of Financial Studies* 1 (Summer 1988), 137–58.
10. Philip R. Perry. "Portfolio Serial Correlation and Nonsynchronous Trading." *Journal of Financial and Quantitative Analysis* 20 (December 1985), 517–23.
11. Myron Scholes and Joseph Williams. "Estimating Betas from Nonsynchronous Data." *Journal of Financial Economics* 5 (December 1977), 309–27.
12. Jay Shanken. "Nonsynchronous Trading and the Covariance-Factor Structure of Returns." *Journal of Finance* 42 (June 1987), 221–31.
13. Hans R. Stoll and Robert E. Whaley. "The Dynamics of Stock Index and Stock Index Futures Prices." Working paper, Owen Graduate School of Management at Vanderbilt University, 1987.