Review of Quantitative Finance and Accounting, 1 (1991): 135-151


# Kalman Filter Estimation for Valuing Nontrading Securities, with Applications to the MMI Cash-Future Spread on October 19 and 20, 1987

GILBERT W. BASSETT, JR.
*Department of Economics, University of Illinois at Chicago, Chicago, IL 60680-2451*

VIRGINIA G. FRANCE
*Department of Finance, University of Illinois at Urbana-Champaign, Champaign, IL 61820*

STANLEY R. PLISKA
*Department of Finance, University of Illinois at Chicago, Chicago, IL 60680-2451*

**Abstract.** The Kalman filter is proposed as a method for estimating the value of nontrading securities during periods when other securities are trading. The method also provides confidence intervals that indicate the degree of uncertainty regarding estimated value. The method is applied to the Major Market Index during the principal days of the 1987 stock market crash. Our results indicate that nonsynchronous trading explains a small but significant portion of the cash-futures spread that prevailed during these days.

**Key words:** Kalman filter, nontrading securities, value estimates

## 1. Introduction

The value of a stock index and the associated cash-futures spread cannot be observed directly when stocks in the index are not trading. To estimate the value of the index requires estimates of the values of the nontrading stocks. Associated with this estimate will be some degree of uncertainty, the amount depending on 1) the proportion of stocks in the index that are not trading, 2) the length of time since the last trade, and 3) the correlation between trading and nontrading stocks. Therefore one seeks a method for estimating the value of an index and the associated measure of uncertainty for situations when discontinuities in trading of the constituent stocks preclude knowing the exact cash value of the index.

This article describes how the Kalman filter can be used for this purpose. The Kalman filter is a well-known recursive procedure for computing an estimate of the state of a linear dynamic system subject to noisy and incomplete observations; for a review, see Harvey (1982). The procedure has been applied to other aspects of finance. Wolff (1987a) used the Kalman filter to estimate time-varying parameters to improve the predictive performance of a class of monetary exchange-rate models. Wolff (1987b) also used the method to identify and measure premia in the pricing of forward exchange rates. Some other applications of the Kalman filter in finance can be found in Brown et al. (1983), Bos and Newbold (1984), Hamilton (1985), and Burmeister and Wall (1986).

The method is applied here to the valuation of the Major Market Index (MMI), which consists of the 20 stocks listed in table 1. The valuation method is important for stock index

*Table 1.* Standard deviations of one-minute price differences, October 16, 1987

| Stock | Symbol | $\sigma$ |
|---|---|---|
| Merck & Co. | MRK | .4234 |
| International Business Machines | IBM | .2053 |
| Phillip Morris | MO | .1702 |
| Dow Chemical | DOW | .2826 |
| Procter & Gamble | PG | .1946 |
| Du Pont | DD | .2423 |
| Johnson & Johnson | JNJ | .1773 |
| Minnesota Mining & Manufacturing | MMM | .1922 |
| General Motors | GM | .1664 |
| Eastman Kodak | EK | .2084 |
| General Electric | GE | .1358 |
| International Paper | IP | .1253 |
| Exxon | XON | .0918 |
| Coca Cola | KO | .1311 |
| Chevron | CHV | .0902 |
| Mobil | MOB | .1132 |
| Sears Roebuck | S | .0889 |
| USX Corporation | X | .0687 |
| AT&T | T | .1007 |
| American Express | AXP | .1122 |

futures because cash settlement is determined by the prices of the stocks in the index. If trading is nearly continuous, any discrepancy between stock and futures prices will be nearly eliminated by arbitrage. Certainty of the cash value of the index, however, does not occur when there are trading gaps. The index value must then be estimated without knowing the exact current values of all constituent stocks

The equilibrium value of the cash-futures spread is a function of several variables. The futures and cash prices should be related to each other by an arbitrage transaction that involves taking opposite positions in the cash and futures instruments and borrowing or lending. The cost of the transaction, and therefore the no-arbitrage bounds on the futures price, is a function of the borrowing and lending opportunity costs of the market participants, the expected dividend flows from the underlying stocks, and the transaction costs in the cash and futures markets. Usually, different types of firms will be involved in long and short arbitrage (see Holden, 1989). In general, there will be upper and lower no-arbitrage bounds on the cash-futures spread and therefore a range for the theoretical basis. Using the expected dividend stream published by the Chicago Board of Trade (CBOT) and bid-ask quotes on T bills, the theoretical basis should have been .60 to .62 index points on October 16th, i.e., .13% of the index value. Transaction costs, such as commissions or bid-ask spreads on the stocks or futures, could substantially widen the no-arbitrage bounds.

Our methods are illustrated for Monday and Tuesday, October 19 and 20, 1987. Minute-by-minute point and confidence-interval estimates for the cash value of the index are proposed for both days. Monday was the day of the market crash, when security values fell by more than 20%. The focus for Monday is estimates of the value of the index during the first two hours, when there were delayed openings of stocks. After the open Monday, there were (with few exceptions) no large gaps in trading; stocks traded almost continuously despite the steep drop in prices. With few trading gaps, the estimate of cash value using the Kalman filter (or any other method designed to account for trading gaps) will necessarily be similar to the cash value estimated from most recent trades.

While Tuesday is not as famous, it presents estimation problems that are more difficult and interesting than Monday. Unlike Monday, trading on Tuesday was not continuous after the open. Stocks in the MMI and the market as a whole opened nearly on time and at levels well above the Monday close. Large sell orders then began to accumulate, and at 10:00 (Chicago time) many stocks ceased trading. By noon, sell orders had become so large that the New York Stock Exchange (NYSE) considered suspending trading. Anticipation of the trading halt led the Chicago Mercantile Exchange to suspend the Standard and Poor's (S&P) 500 futures contract; the MMI futures contract, however, stayed open throughout the day.

The large price swings in the MMI futures and cash values for Monday and Tuesday are depicted in figure 1; figure 2 depicts the associated cash-futures basis. The discontinuous trading that occurred for a sample of stocks is illustrated by figure 3. Notice that for nearly two hours, midday Tuesday, there were no trades for some of the stocks. A key problem to be addressed is how to value each of the stocks during this period.[1]

The large spreads between the stock and futures prices shown in figures 1 and 2 are based on the last trade estimate of the cash value of the futures contract. For the reasons discussed above, this might not be a realistic estimate of the actual cash value of the index. Consequently, the large spread that is reported may be an artifact. Better estimates of the cash index can possibly provide a more realistic picture of the difference between stock and futures values.

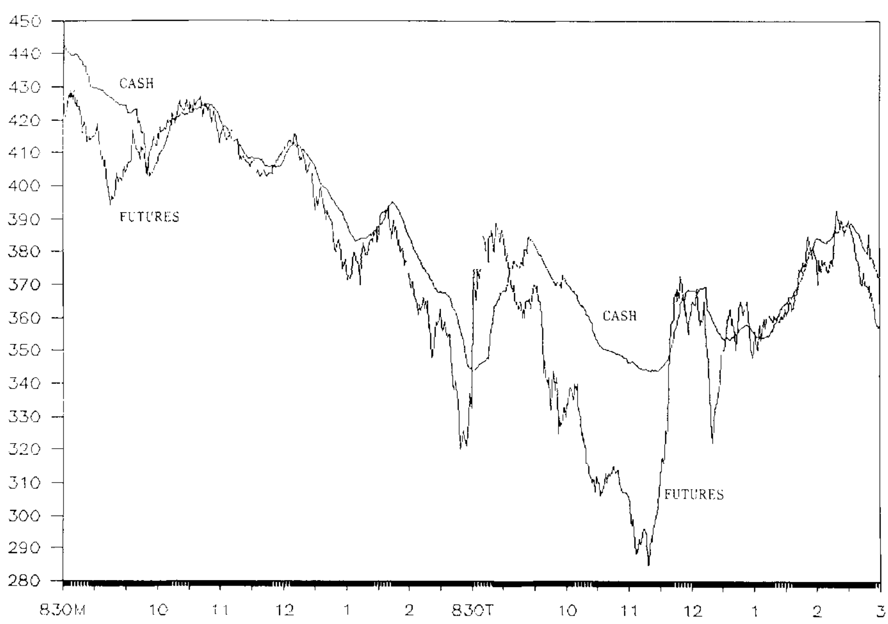


*Figure 1.* MMI futures and cash indices, October 19 and 20, 1987.

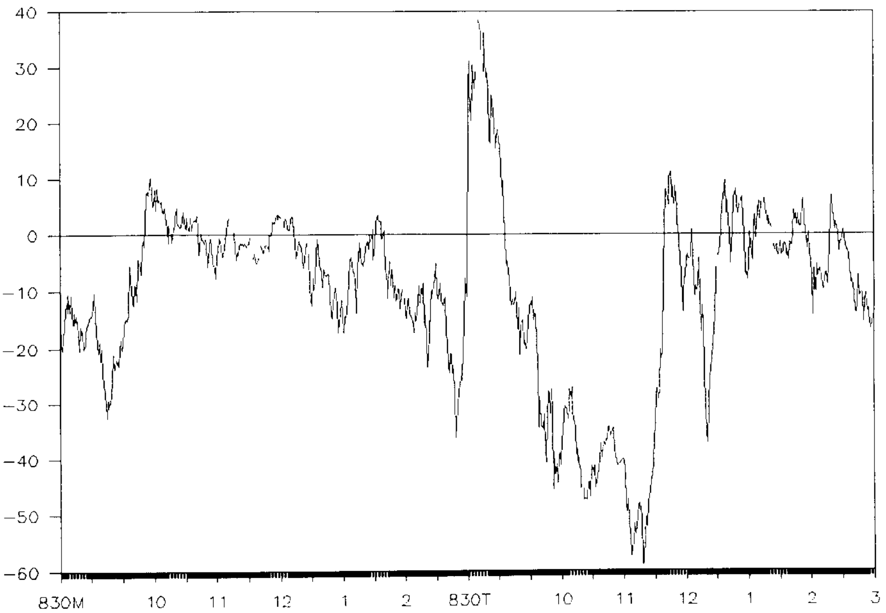


*Figure 2.* Cash futures basis, October 19 and 20, 1987.

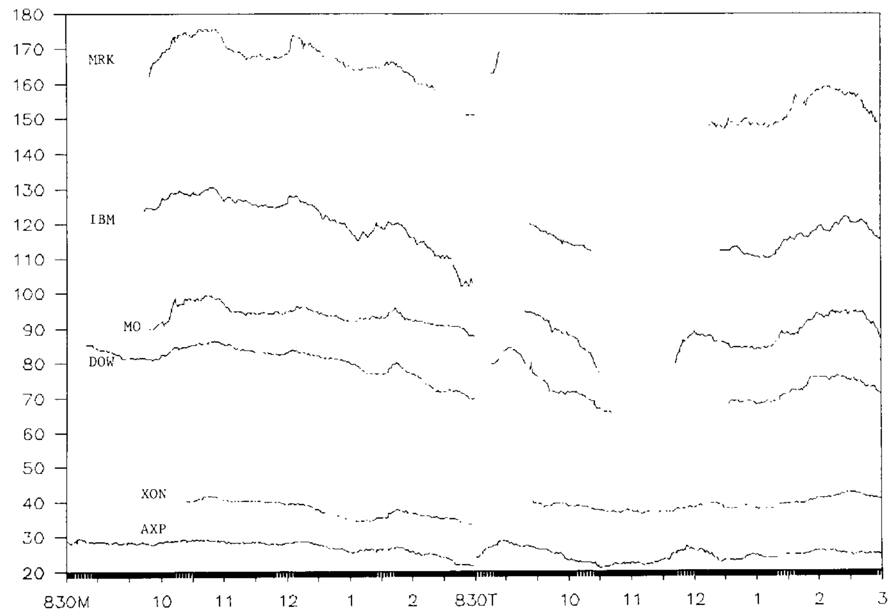


*Figure 3.* Selected stock prices, October 19 and 20, 1987.

Our results complement previous studies that have attempted to estimate the value of stock index futures on the days around the crash; see Edwards (1989) for a review of the government reports on the crash, Harris (1989) for analysis of the S&P 500 contract, Blume et al. (1989), who examine pricing anomalies on Monday and Tuesday, and Bassett, France and Pliska (1989) for analysis of the MMI cash-futures spread on Monday.

The emphasis of previous work has been on deriving point estimates of the cash value of the stock index contracts during the times on Monday and Tuesday when stocks were not trading. Cash values of nontrading stocks were estimated using nearby prices in both the past and future. This contrasts with the Kalman filter methods considered here. These estimates can be implemented in real time because they only require prior price information. The method also provides convenient interval estimates for representing the uncertainty of the cash index value when there are nontrading stocks.

The Kalman filter is described in the next section. The parameter estimates needed to compute estimates of the cash values are discussed in section 3. The application of the model to MMI data on Monday and Tuesday is presented in section 4. Conclusions, other applications of the Kalman filter approach, and extensions of the analysis are discussed in the last section.

## 2. The Kalman filter model

We use a discrete time model where each period ($1 \leq t \leq 390$) corresponds to one minute during the trading day on the NYSE. Let $X_t \epsilon R^{20}$ be a column vector representing the true values of the 20 MMI stocks at time $t$. We assume

$$X_{t+1} = X_t + W_t$$

for all $t$, where $\{W_t\}$ is an *iid* sequence of random vectors with $E[W_t] = 0$ and covariance matrix, $\text{Var}(W_t) = Q$.[2]

Suppose $k$ ($0 \leq k \leq 20$) stocks, indexed by $j(1), j(2), \ldots, j(k)$, trade at time $t$ at prices given by $Y_t(1)$, $Y_t(2)$, ... $Y_t(k)$. The column vector $Y_t \epsilon R^k$ denotes these prices. Let $B_t$ denote the $k \times 20$ matrix that indicates which stocks trade at time $t$; that is,

$$B_t(i, j) = \begin{cases} 1, & j = j(i) \quad i = 1, \ldots, k \\ 0, & j \neq j(i). \end{cases}$$

We assume the vector $Y_t$ of observed prices is related to $X_t$ by

$$Y_t = B_t X_t + Z_t,$$

where $\{Z_t\}$ is a sequence of independent random vectors with $E[Z_t] = 0$ and $\text{Var}(Z_t) = H$. This specification makes observed prices, $Y_t$, a noisy version of $X_t$, the values of stocks. The discrepancy between $X_t$ and $Y_t$ exists because of the bid-ask spread and the discreteness of prices due to the 1/8 tick size.[3]

Given a history $Y_1, Y_2, \ldots, Y_t$ of price observations, the estimation problem is to compute an estimate $\hat{X}_t$ of the vector $X_t$ of stock values.[4] The Kalman filter algorithm is ideally suited for this purpose, since it is simple and recursive and can be implemented in real time. Detailed descriptions of the approach may be found in Davis (1977) or Krishnan (1984).

To illustrate how one iteration of the algorithm works, let $\hat{X}_{t-1}$ and $C_{t-1}$, the covariance matrix associated with $\hat{X}_{t-1}$, be given. Since $\hat{X}_{t-1}$ is also a forecast of $X_t$, the "forecast error" after observing $Y_t$ is the column vector,

$$V_t = Y_t - B_t\hat{X}_{t-1}.$$

Associated with $V_t$ is the $(k \times k)$ covariance matrix

$$F_t = B_t(C_{t-1} + Q)B_t' + H_t.$$

The estimate of $X_t$ is then given by

$$\hat{X}_t = \hat{X}_{t-1} + (C_{t-1} + Q)B_t'F_t^{-1}V_t.$$

The covariance matrix associated with $\hat{X}_t$ is given by

$$C_t = C_{t-1} + Q^{-1}(C_{t-1} + Q)B_t'F_t^{-1}B_t(C_{t-1} + Q).$$

Given the estimated values of the stocks, the estimate of the MMI cash value for period $t$ is given by,

$$I_t = (\hat{X}_t'1)d^{-1},$$

where $\mathbf{1} \epsilon R^{20}$ is a column vector of ones and the scalar $d$ is the MMI divisor. The variance associated with the estimate $I_t$ is given by $(\mathbf{1}'C_t\mathbf{1})d^{-2}$.

## 3. Parameter estimation

This section discusses the four quantities—$Q = \text{Var}(X_t)$, $H = \text{Var}(Z_t)$, $\hat{X}_0$, and $C_0$—that are needed to implement the Kalman filter algorithm.

The data base consists of minute-by-minute prices for each of the 20 MMI securities and the MMI futures contract.[5] The data are reported from 8:30 to 3:00, Chicago time, for three days: Friday, October 16; Monday, October 19; and Tuesday, October 20, 1987.

The estimate of the covariance matrix $Q$ is based on the empirical covariances derived from Friday data.[6] The estimated covariances used for Monday and Tuesday are thus based on the most recently observed covariances between the MMI stocks. Further, they are based on a day when prices also fell, though not by as much as on Monday. (Recall that the Dow Jones Index (DJI) fell by 108 points on Friday.) The standard deviations of first differences in the minute-by-minute price series are listed in table 1; the full $20 \times 20$ covariance matrix is available upon request.

The covariance matrix $H$ was chosen based on consideration of the underlying market structure as well as our finding that its properties are analogous to an exponential smoothing parameter. In particular, $H$ is taken to be diagonal with elements equal to .005. This variance corresponds to a standard deviation of .07, a number that is consistent with the bid-ask spread and the 1/8 tick size.

To help illustrate how the choices of $Q$ and $H$ affect the performance of the Kalman filter algorithm, consider the following simple example. Suppose there are only two stocks with $\hat{X}_0 = (10, 100)$ and

$$Q = C_0 = \begin{bmatrix} .01 & .1r \\ .1r & 1.0 \end{bmatrix}, \tag{1}$$

where $Q$ depends on $r$, the correlation coefficient between the two stocks. Suppose that in period 1, only the first stock trades at the price 9.75, a 2.5% drop.

Table 2 shows the period 1 estimates for both stocks as a function of alternative values for $r$ and $H$. The estimate for the first stock depends only on $H$, which here is a scalar. The estimate for the second stock, the one that did not trade, is also shown. This estimate depends on both $r$ and $H$. In all cases, the smaller the value of $H$, and thus the greater the accuracy of the observations, the greater the difference between the period 0 and period 1 estimates. Thus the smaller values of $H$ make the estimates more sensitive to recent observations. The dependence of the estimated value of the second stock with respect to the correlation coefficient $r$ also accords with intuition: the larger the value of $r$, the more the estimate shifts in the direction of change for the first stock.

We now turn to the selection of $\hat{X}_0$ and $C_0$, the values for the initial minute of trading during a trading day. While one would normally choose $\hat{X}_0$ to coincide with the estimate for the last minute of the preceding day, it is not clear how to choose the covariance matrix, $C_0$. The gap of 17.5 hours (65.5 over a weekend) since the most recent trades serves to increase one's uncertainty about opening prices. Hence it may be inappropriate to set $C_0$ equal to the value of $C_{390}$ from the preceding day.

The approach adopted here is as follows. Assume that the clock-time gap between the close on day $d$ and the opening on the next is equivalent to a gap of *equivalent nontrading minutes*, denoted by $N(d)$. The algorithm will run continuously from one day to the next,

*Table 2.* Illustration of parameter insensitivity

| | | $X_1(2)$ as a function of $r$ | | | | | |
|---|---|---|---|---|---|---|---|
| $H$ | $X_1(1)$ | 0.9 | 0.5 | 0.1 | −0.1 | −0.5 | −0.9 |
| .5 | 9.99 | 99.91 | 99.95 | 99.99 | 100.01 | 100.05 | 100.09 |
| .1 | 9.96 | 99.63 | 99.79 | 99.96 | 100.04 | 100.21 | 100.38 |
| .05 | 9.93 | 99.36 | 99.64 | 99.93 | 100.07 | 100.36 | 100.64 |
| .01 | 9.83 | 98.50 | 99.17 | 99.83 | 100.17 | 100.83 | 101.50 |
| .005 | 9.80 | 98.20 | 99.00 | 99.80 | 100.20 | 101.00 | 101.80 |
| .0001 | 9.75 | 97.76 | 98.76 | 99.75 | 100.25 | 101.24 | 102.24 |

with intervals of 390 trading minutes separated by nontrading intervals of length $N(d)$ minutes. During each minute of a nontrading interval, the vector $Y_t$ will be null, so the estimate corresponding to the opening of the market will remain unchanged. However, the covariance matrix $C_t$ will change as the estimates during the nontrading minutes become less certain.

For example, the specification $N(d) = 0$ means that there is no effective gap in the evolution of the value process when markets are closed. It would be as if values at the open are connected directly to the previous close, with no intervening gap in trading. The effect is to collapse the distribution of $X_0$, conditional on observed prices from the previous close, onto the closing prices from the previous day. This contrasts with the opposite extreme in which $N(d)$ is the literal number of minutes between the close and subsequent open. There will then be considerable dispersion at the opening before stocks begin trading.[7]

To illustrate the effect of variations in the overnight parameter, consider the above example and let the stock index be the mean of the two prices. Suppose $r = 0$ with day $d = 1$ concluding with $X_{390}(1) = (10{,}100)$. The estimate for the open on $d = 2$ will be $I_0(2) = 55 = \frac{1}{2}(10 + 100)$, the same as the value at the close on day $d = 1$. The variance of the estimate depends on $N(1)$ and in this simple case can be expressed as $N(1)*(1.01/4)$. This example shows the extent to which the uncertainty regarding opening values is influenced by the length of the overnight trading gap.

To conclude this discussion of implementation of the Kalman filter, the initial values for $X_0$ and $C_0$ on Friday were arbitrarily based on Friday data. The algorithm for Friday, October 16, was implemented with $X_0$ equal to the actual opening trades on Friday and with $C_0$ equal to $Q$. It was found that the estimates for the closing values on Friday were not sensitive to these initial specifications.[8]

## 4. Results

When a stock is trading almost continuously, the Kalman filter estimate of the stock's value differs only slightly from the transaction price. This difference, which reflects the bid-ask spread and the discreteness of price movements, is small and cannot be observed on the accompanying graphs.

If a stock does not trade during an interval of time, the Kalman filter produces a point estimate that is based on 1) the stock's previous price, 2) the intervening movement of the trading stocks, and 3) the estimated covariances between prices. This is illustrated in figure 4 for IBM on October 20. IBM opened at 9:20 and traded until about 10:20. Trading then stopped until about 12:25. Note that the Kalman filter estimates during the intervals 8:30–9:20 and 10:20–12:25 move in a fashion that is consistent with the futures and stocks (see figures 1–3). It also should be noted that the estimates just after 10:20 are similar to the transaction prices just before 10:20, whereas the estimates just before 12:25 are much higher than the transaction prices just after 12:25. The discontinuity at 12:25 occurred because the uncertainty about the estimates increases as time elapses since the last trade. This is seen in the confidence-interval estimates for IBM during the gap in trading. Figure 5 shows how the 95% confidence interval balloons outward as the time since the last trade increases.[9]

The computation of the Kalman filter estimate of the MMI index value begins on Monday the 19th and continues through the end of the 20th. The first set of estimates is based on

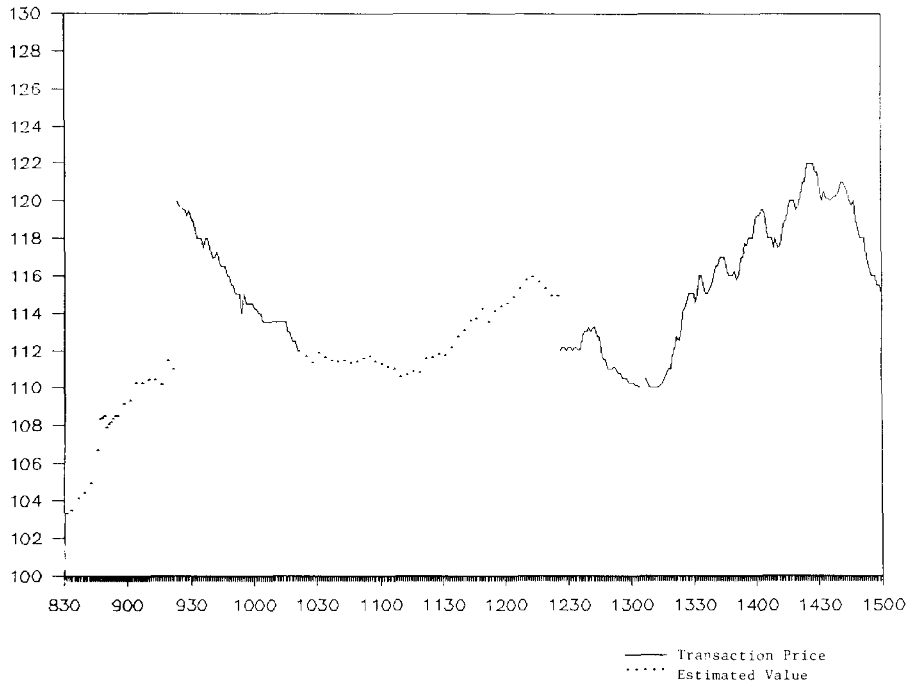


*Figure 4.* IBM prices and estimated value, Tuesday, October 20, 1987.

$N(1) = N(2) = 0$, that is, no dummy minutes of trading were inserted between the close on one day and the subsequent open. Hence, the Kalman filter estimate of the Monday opening does not substantially differ from the Friday closing prices, and the size of the confidence interval is nearly zero.

Figure 6 compares the conventional cash basis and Kalman filter basis for Monday and Tuesday. Figure 7 shows the 95% confidence interval for the Kalman filter estimates on Monday and Tuesday.

Many stocks in the index opened late on Monday morning. Until a stock opens, the conventional last-trade index is computed using the previous day's close. Since those stocks that were trading showed declines on Monday, the conventional cash index was biased upwards by the delayed openings. This is one of the reasons for the large negative cash-futures basis observed on Monday; see Bassett, France, and Pliska (1989).

For the Kalman filter estimate of the index, missing stock prices are not assumed to have remained constant, as is implicitly assumed when using a last-trade-based estimate of cash values. Thus, Kalman filter estimates of the index for Monday morning are lower than the conventional cash index. The absolute value of the basis is reduced on Monday morning using the Kalman filter estimates of the cash value of the index.

After all the MMI stocks opened on the 19th, there were only brief gaps in trading of stocks. As a result, the Kalman filter estimate and conventional estimates of the index do not differ greatly, and the confidence intervals are small.

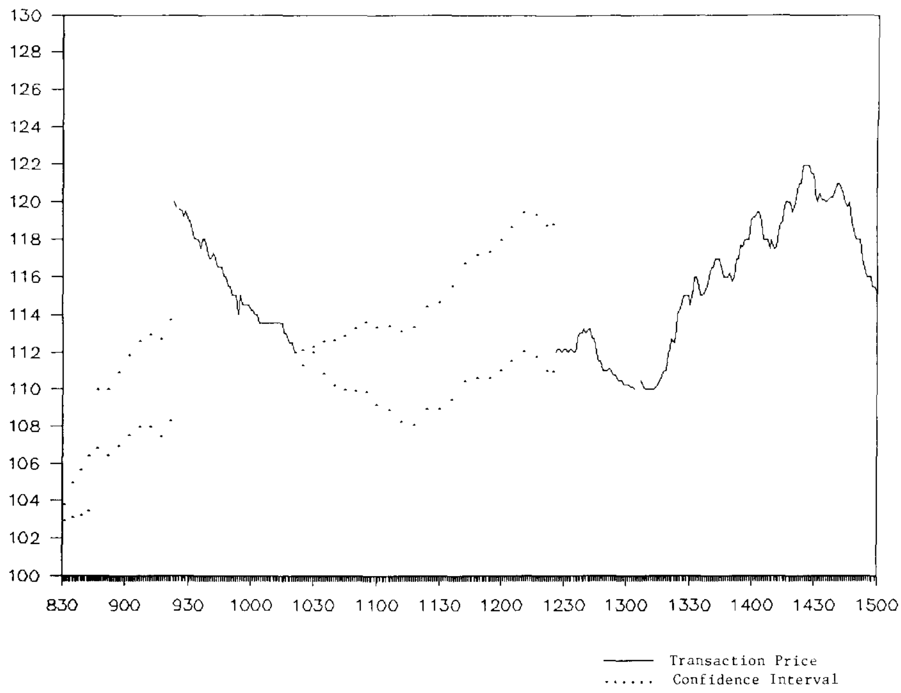


*Figure 5.* IBM confidence intervals, Tuesday, October 20, 1987.

On Tuesday morning there was a brief rally, and some stocks opened late. The conventional basis is now biased upwards by the late-opening stocks, and the basis obtained from the Kalman filter estimates is again closer to zero.

After about 9:15 the rally in both the futures and stock markets came to an abrupt halt, and prices began to fall. Large sell orders led to halts in trading of many stocks in the MMI index. For example, IBM stopped trading at 10:20 with a transaction at $112, and did not resume trading until about 12:25 with a transaction at about the same value. Linear interpolation over this gap leads to the conclusion that the value of IBM was almost constant during the two-hour period. The Kalman filter estimates, by contrast, take into account what is known of the behavior of other prices at the time. During the two-hour gap in trading, the estimated value of IBM drops, then increases, and then drops again. Although the magnitudes of these fluctuations are not as great, they do mimic the behavior of the futures price over the same period.

During this large gap in trading, the confidence interval for the value of IBM widens to about 8% of its estimated value. The other three most heavily weighted stocks, Dow Chemical, Philip Morris, and Merck, also have substantial trading gaps during this period. These four stocks make up more than a third of the index value. The large degree of uncertainty about the value of the stocks is reflected in the confidence intervals for the index shown in figure 7.

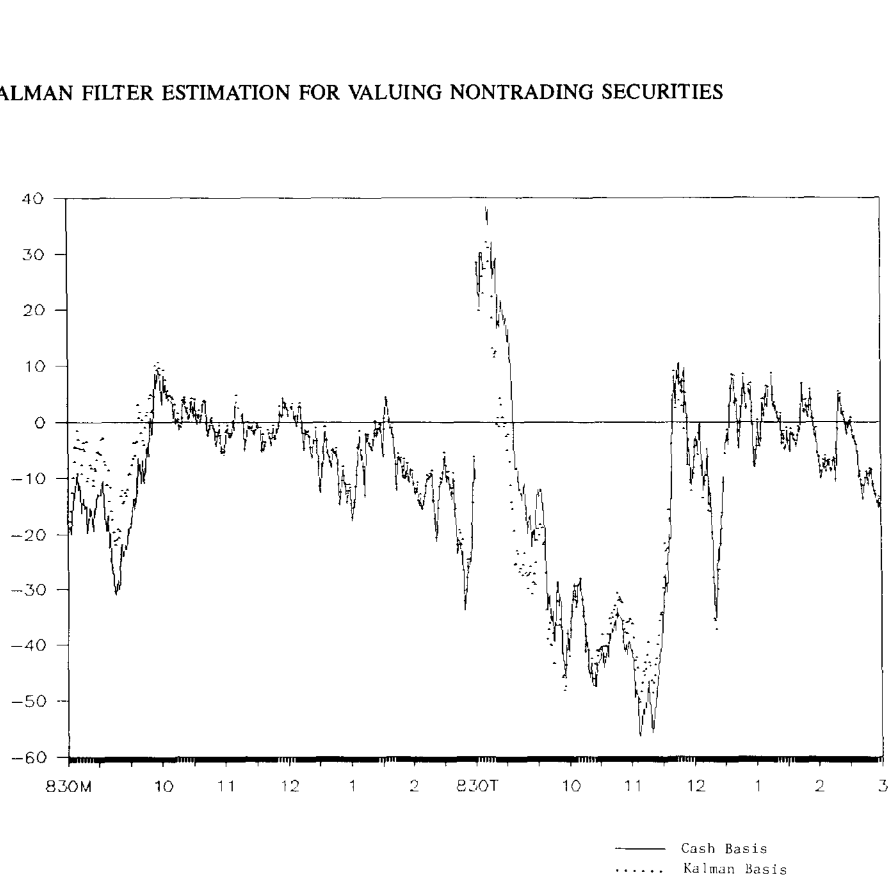


*Figure 6.* Cash basis and Kalman basis, October 19 and 20, 1987.

As discussed in section 3, there is more uncertainty about opening stock values (given the previous closing prices) than there is regarding minute-to-minute values during a trading day. Hence, the confidence interval for the Tuesday open, which is based on $N = 0$, is unrealistically small. This is corrected by inserting dummy trading gaps between the Monday close and the subsequent open.

Figures 8, 9, and 10 show the 95% confidence intervals for Tuesday when one, two, and 17.5 hours of dummy trading are inserted between Monday and Tuesday. These figures end at 1:00, because afterwards all the stocks trade frequently. The figures show that an increase in dummy trading time leads to a widening of confidence intervals with an attenuated effect following transactions. Inserting a gap of 17.5 hours causes the confidence intervals to "capture" the opening futures price, although the price departs from the confidence interval for much of the rest of the morning.[10]

## 5. Discussion

The Kalman filter reduced the absolute value of the cash-futures spread on October 19 and 20, though sometimes by less than 50%. This is similar to the results obtained in previous studies of the crash.

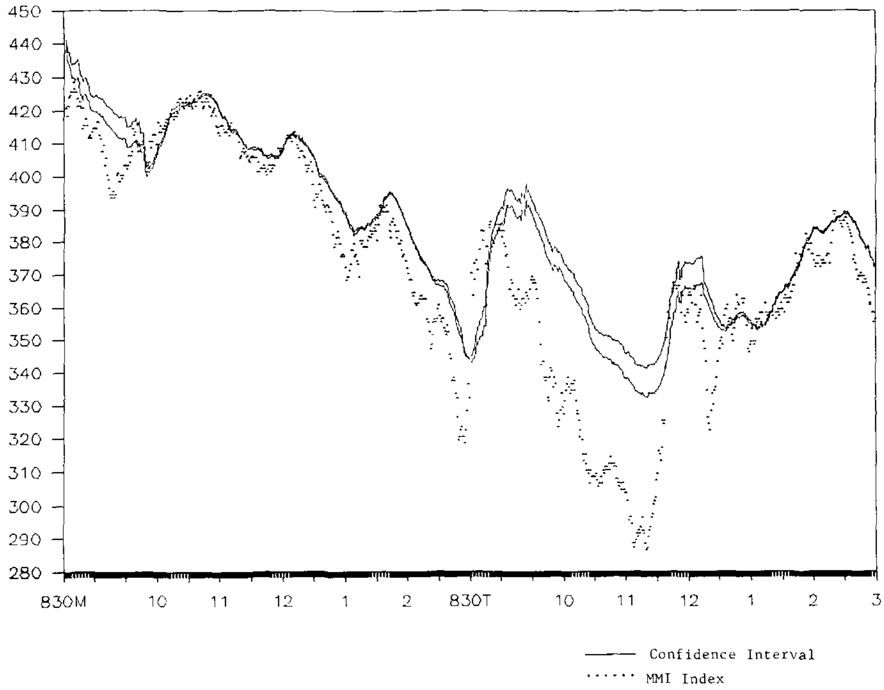


*Figure 7.* Confidence intervals for MMI index values, October 19 and 20, 1987 (overnight gap = 0 hours).

Harris (1989) estimated the cash value of the S&P 500 index using five-minute intervals for the period October 12th through 23rd, 1987. The estimate is based on a weighted least squares estimate of an underlying factor's return. The underlying factor is assumed to affect all stock returns by the same amount in a given period. Since the S&P 500 index is value-weighted (recall that the MMI is based on the sum of stock prices), he uses a weighted least squares regression with the weights set to the value weight of each stock in the index. The estimates of the factor can then be interpreted as estimates of the percent change in the market value of the index. It provides an estimate of cash value that attempts to control for nonsynchronous trading of stocks.

A rolling regression spanning a subinterval is used by Harris to estimate cash value where the size of the regression increases with the number of observations. Two sets of regressions are used, one centered on the period to be estimated (*perfect foresight* regressions) and one that ends with the period to be estimated (*current information* regressions). This contrasts with the Kalman filter that at each period is based on the covariance matrix of one-period returns and a sufficient statistic for the entire history of observations. The Kalman filter was designed for real-time applications, and its computational complexity is considerably less than the technique used by Harris.

The results of Harris (1989) are broadly similar to ours. Nonsynchronous trading explains some but not all of the variation in the basis, and the adjusted index tracks the futures price much better at the open.

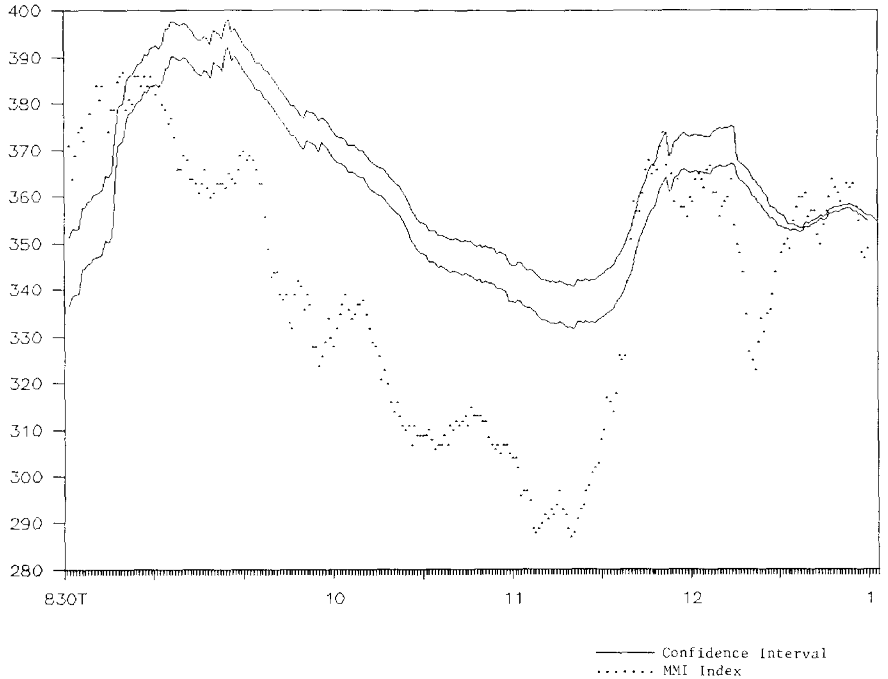


*Figure 8.* Confidence intervals for MMI index values, 8:30–1:00, October 20, 1987 (overnight gap = 1 hour).

Blume, MacKinlay, and Terker (1989) studied relationships between order imbalances and stock price movements on October 19 and 20. Their objective was to understand the linkages between the S&P 500 futures price, the cash prices of NYSE stocks in the S&P index, and the cash prices of NYSE stocks not in the S&P index. They recognized the implications of nonsynchronous trading and the fact that "the published S&P index may underestimate losses in a falling market." They noted, however, that returns of NYSE stocks differed according to whether the stock was in the S&P index, thereby suggesting that futures-related trading may have led to differences in the volume and order patterns for index versus non-index stocks.

To control for the impact of nonsynchronous trading, Blume, MacKinley, and Terker (1989) developed a simple procedure for estimating the cash value of an index when constitutent stocks are not trading. The values of nontrading stocks at a point in time were estimated using the change in the total market value of those stocks that had traded in the preceding 15 minutes; this is similar to the *backward extrapolation* method in Bassett, France, and Pliska (1989). The procedure decreased but did not eliminate the spread during the first 90 minutes of trading on October 19. This too is similar to our finding using the Kalman filter and suggests that the large spreads on October 19 and 20 were not primarily due to nonsynchronous trading.

We originally expected the Kalman filter estimates to produce greater reductions in the spread. The method uses observed transactions, some of which (for example, American

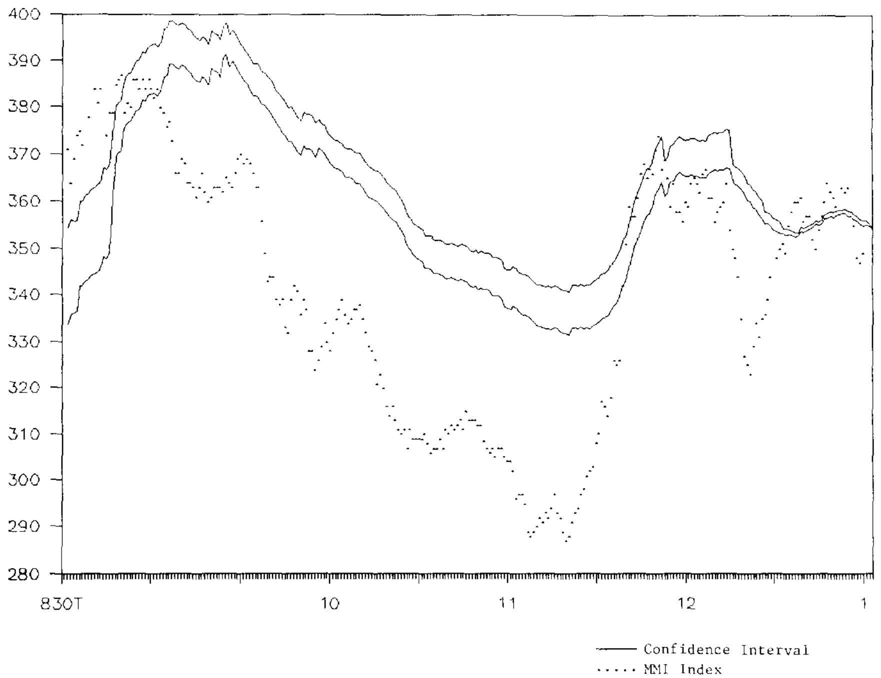


*Figure 9.* Confidence intervals for MMI index values, 8:30–1:00, October 20, 1987 (overnight gap = 2 hours).

Express) were frequent and highly correlated with the futures price. We thus expected the Kalman filter estimates of the index to be more highly correlated with the futures price.

Compared to the naive estimates of cash values based on the last trade, the Kalman filter estimates resulted in a reduction of about 50% in the value of the spread.[11] The failure to reduce the cash-futures spread even further may be due to a combination of the following. First, our estimates of input parameters (especially the covariance matrix) is based on a small sample and might not accurately reflect parameter values on the day of the crash. Second, our application of the Kalman filter ignored nonprice information. There may be selection bias from sell-order backlogs that tend to cause trading gaps in those stocks having the largest decrease in value. Third, statistical assumptions (e.g., normality) that justify the Kalman filter may fail to hold. Better estimates of the cash value of nontrading stocks would be obtained using estimation methods based on more realistic statistical assumptions. Finally, it is possible that there was a considerable difference in stock and futures values at about the same point in time. This explanation goes beyond the inability to execute arbitrage transactions between close substitutes. It says that the chaotic trading conditions were so extreme and the looming financial crisis so great as to create the possibility of default on existing contracts; futures and stocks, therefore, would no longer be nearly perfect substitutes.[12]

We suspect that the explanation for the large cash-futures spread even after using the Kalman filter lies with order imbalances, which caused stocks to cease trading or fail to

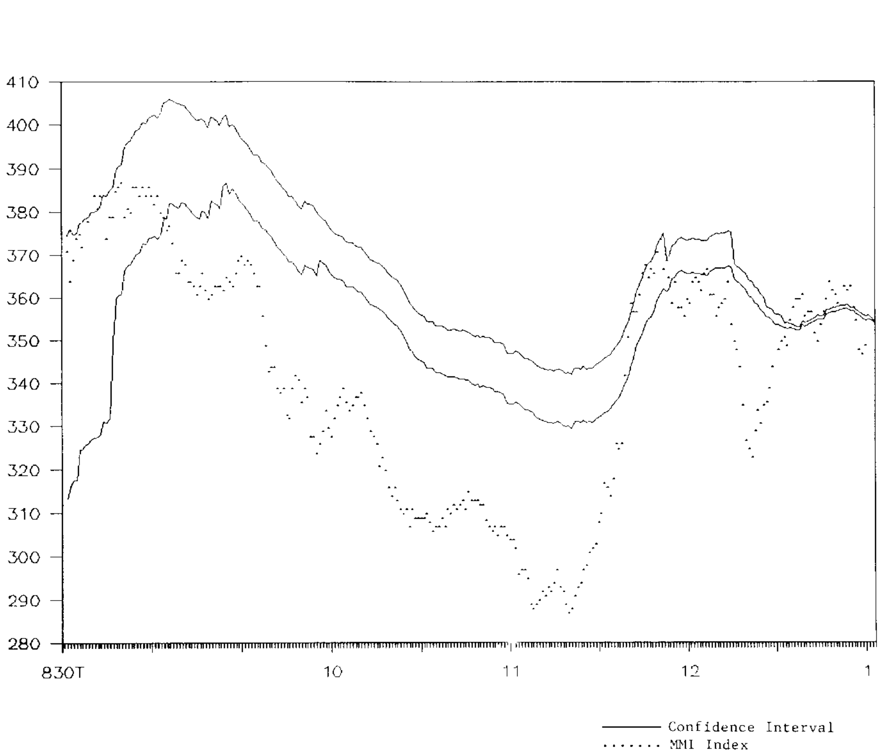


*Figure 10.* Confidence intervals for MMI index values, 8:30–1:00, October 20, 1987 (overnight gap = 17 hours).

open on time. In this case, an estimation method using only trading stocks will always yield a biased estimate for the value of nontrading stocks and hence of the value of the index itself. This selection bias will persist, whether one uses a simple estimation procedure or a more sophisticated procedure such as the Kalman filter.

An important added feature of the Kalman filter is its associated uncertainty measure for the current index value. This uncertainty is especially relevant for assessing the large basis that occurred on the crash days when there were gaps in trading. Moreover, it is of more general use for assessing the uncertainty of opening values based on previous trades and the "effective" trading gap when markets are closed.

The confidence intervals can also be used to assess the risks in apparent arbitrage opportunities. One might judge whether there is a legitimate arbitrage opportunity by comparing not only the futures price with an estimate of cash value, but, in addition, the futures price to a confidence band reflecting uncertainty of the cash value.

The method can be used to estimate the NYSE open for indices or individual stocks. It would allow a specialist, for instance, to incorporate data from overnight transactions in American Depository Receipts (ADRs), or prices observed on the London exchange. Alternatively, the method can be used to estimate the S&P 500 cash index before all 500 stocks have opened (perhaps to establish a cash-futures arbitrage position).[13]

Perhaps a more important application of the estimation method is futures contracts on international indices. Morgan Stanley's Europe-Australia-Far East (EAFE) index represents the value of stocks that trade on exchanges from Europe to Australia. At any time, many of the prices which make up the index are stale, because the relevant exchanges are open different hours. Estimating the EAFE index value is qualitatively identical with the procedure investigated here. The Kalman filter would provide the exchange with a method for posting index values when components of the index are not trading.

## Notes

* Portions of this research were supported by a grant from the Educational Research Foundation of the Chicago Board of Trade. Professor France's research was in part supported by the Investors in Business Education, University of Illinois at Urbana-Champaign. We would like to thank participants in seminars at the University of Illinois at Chicago and the University of Illinois at Urbana-Champaign for their comments. We would also like to thank Jen Chin Chen for research assistance, Ted Doukas for providing the MMI futures data, and a referee for useful comments.

1. A related question, which will not be addressed here, is the interrelationship between world security markets; for a recent study, see King and Wadhwani (1990).
2. When constructing confidence intervals, it is further assumed that $X_{t+1}$ is normally distributed.
3. It is reasonable to regard the value $X_t$ as falling somewhere between the bid and asked prices. If the transaction prices fluctuate between the bid and asked prices, then the transaction prices will fluctuate around the underlying value $X_t$ without necessarily coinciding with it.
4. An implicit assumption of the model is that the value of a stock does not depend on whether or not it trades; that is, $X_t$ does not depend on $B_t$. Under this assumption, the estimates for nontrading stocks only use information on prices of trading stocks. For its application to the days of the crash, this implicit assumption fails for some time periods. The stocks that had large gaps did not trade precisely because of large order imbalances. Conversely, the stocks that did trade tended to be those that were least affected by the chaotic trading conditions. As a result, the Kalman filter estimates of the change in value are biased toward zero, and the dispersion of $X_t$ depends on $B_t$. Representing this in the model and accounting for selection bias in the estimates is a topic for future research.
5. The stock data were extracted from transaction data, which, in turn, were obtained from the Francis Emory Fitch Company. In order to get the most recent price for the cash stocks as of the start of each minute, we used the last price in the previous minute. If there was no trade during that minute, the price was treated as missing. Thus, our data set includes the transaction price that would be used to construct a last-trade cash index, which allowed us to check our data against the reported cash index. We then compare this with the futures transaction that immediately follows, that is, the first futures price in each minute. Transaction data on the MMI futures contract and a last-trade cash index were provided by the Chicago Board of Trade. Since the futures is extremely liquid, we are able to narrow the gap between when we observe the cash and the futures to a few seconds, though the cash data are always very slightly older.
6. A continuous minute-by-minute series was constructed by filling in any trading gaps with the last trade price. Covariances were estimated from prices between 8:47, the time when the last MMI stock opened, until 3:00. The 373 one-minute price differences were then used to estimate the covariance matrix for the 20 MMI stocks.
7. There have been several studies estimating price variability during the times when markets are closed; see, for example, Fama (1965), French and Roll (1986), and Oldfield and Rogalski (1980).
8. An additional complication was Eastman Kodak's 3-for-2 stock split that occurred betwen Monday and Tuesday. This led to a change in the MMI divisor from 3.18322 on Monday to 3.12165 on Tuesday. Since Eastman Kodak did not open until 9:55 on Tuesday, the customary cash index before 9:55 on Tuesday was computed using the last Monday trade of Eastman Kodak. For Tuesday we replaced the Eastman Kodak component of the MMI by two thirds of its closing Monday value and used the new divisor from 8:30 onward.
9. The confidence interval is $\pm 1.96$ times the estimated standard error, which is the square root of the diagonal element of $C_t$ corresponding to IBM.
10. The length of the confidence interval is a linear function of $Q$. The futures prices may be outside the interval because our estimate for $Q$, which is based on Friday trading, underestimates the variances for MMI stocks on Monday and Tuesday. Also recall the reasons given above (footnote 4) for expecting the confidence intervals to be conservative and to underestimate the degree of uncertainty.

11. The tape ran late during the crash, sometimes for several hours. Our data are identified by the time the transaction occurred, not the time the trade was reported on the tape. Thus, we avoid some of the worst "stale prices." For a measure of the true value of the index, we have the correct data: the actual transaction prices at the actual transaction times. This information, however, was not available to traders watching the tape and hence could not have been used to know about arbitrage opportunities. Thus, the question remains whether an arbitrageur would have known that a transaction could have been executed at a given price.
12. For example, investors may have considered futures more likely to default on account of their greater leverage and the large margin shortfalls due to the crash on Tuesday morning; see Bernanke (1990).
13. The technique can also be applied generally to estimate an index based on illiquid or infrequently traded assets. For instance, Case and Shiller (1989) compute an index based on a 16-year sample of repeat sales of homes, where transactions can occur several years apart at irregular intervals. Their index is based on the returns between first and second sale. The Kalman filter would allow estimation of the value of individual houses as well as of the whole index.

## References


Bassett, G.W., V.G. France, and S.R. Pliska, "The MMI Cash Futures Spread on October 19, 1987." *The Review of Futures Markets* 8(1), 118–138 (1989).

Bernanke, B.E., "Clearing and Settlement During the Crash." *The Review of Financial Studies* 3(1), 133–151 (1990).

Blume, M.E., A.C. MacKinlay, and B. Terker, "Order Imbalances and Stock Price Movements on October 19 and 20, 1987." *Journal of Finance* 44, 827–848 (1989).

Bos, T. and P. Newbold, "An Empirical Investigation of the Possibility of Stochastic Systematic Risk in the Market Model." *Journal of Business* 57(1), 35–41 (1984).

Brown, P., A. Kleidon, and T. Marsh, "New Evidence on the Nature of Size-Related Anomalies in Stock Prices." *Journal of Financial Economics* 12(1), 105–127 (1983).

Burmeister, E. and K. Wall, "The Arbitrage Pricing Theory and Macroeconomic Factor Measures." *Financial Review* 21(1), 1–20 (1986).

Case, K.E. and R.J. Shiller, "The Efficiency of the Market for Single-Family Homes." *American Economic Review* 79, 125–137 (March 1989).

Davis, M.A. *Linear Estimation and Stochastic Control*. New York: Academic Press, 1982.

Edwards, F.R., "Studies of the 1987 Stock Market Crash: Review and Appraisal." Working Paper No. 168, Center for the Study of Futures Markets, Columbia University, 1989.

Fama, E., "The Behavior of Stock Market Prices." *Journal of Business* 38, 24–105 (1965).

French, K.R. and R. Roll, "Stock Return Variances, The Arival of Information and the Reaction of Traders." *Journal of Financial Economics* 17, 5–26 (1986).

Hamilton, J.D., "Uncovering Financial Market Expectations of Inflation." *Journal of Political Economy* 93(6), 1224–1241 (1985).

Harris, L., "The October 1987 S&P 500 Stock Futures Basis," *Journal of Finance* 44, 77–100 (1989).

Harvey, A.C., "The Kalman Filter and Its Applications in Econometrics and Time Series Analysis," *Methods of Operations Research* 44, 3–18 (1982).

Holden, C., "A Theory of Arbitrage in Financial Market Equilibrium." Working paper, Anderson Graduate School of Management, University of California at Los Angeles, 1989.

King, M.A. and Sushil Wadhwani, "Transmission of Volatility between Stock Markets." *The Review of Financial Studies* 3(1), 5–33 (1990).

Krishnan, V., *Nonlinear Filtering and Smoothing: An Introduction to Martingales, Stochastic Integrals, and Estimation*. New York: John Wiley and Sons, 1984.

Oldfield, G. and R.J. Rogalski, "A Theory of Common Stock Returns Over Trading and Non-Trading Periods." *Journal of Finance* 35, 729–751 (1980).

Wolff, C.C.P., "Time-Varying Parameters and the Out-of-Sample Forecasting Performance of Structural Exchange Rate Models." *Journal of Business and Economic Statistics* 5(1), 87–97 (1987a).

Wolff, C.C.P., "Forward Exchange Rates, Expected Spot Rates, and Premia: A Signal-Extraction Approach." *The Journal of Finance* 42(2), 395–406 (1987b).