New England Economic Review

Federal Reserve Bank of Boston

March/April 1993

Peter Fortune — *Stock Market Crashes: What Have We Learned from October 1987?*

Katharine L. Bradbury — *Equity in School Finance: State Aid to Schools in New England*

Richard W. Kopcke — *Forecasting Investment with Models and Surveys of Capital Spending*

The **New England Economic Review** is produced in the Research Department and is edited by Joan T. Poskanzer. Heather Schofield coordinates the design and production. The views expressed are those of the individual authors, and do not necessarily reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System. The authors will be glad to receive comments on their articles. The Review is available without charge. Requests to be placed on the mailing list or for additional copies should be sent to the Research Library–D, Federal Reserve Bank of Boston, P.O. Box 2076, Boston, MA 02106-2076. Articles may be reprinted if the source is credited. Please send copies of the reprinted materials to the editor.

Contents

Contents

Stock Market Crashes: What Have We Learned from October 1987?

Peter Fortune

Professor of Economics at Tufts University. This paper was written while the author was a Visiting Scholar at the Federal Reserve Bank of Boston. *He is grateful to David Zanger for research assistance, to Stephen Balter of Merrill Lynch for aid in this project, and to Richard Kopcke for constructive comments.*

This is the third article in a trilogy assessing the performance of the U.S. stock market. The first paper examined the historical record of volatility in the markets for bonds and common stocks, concluding that the volatility of the stock market has not increased in recent years but that bond markets are more volatile now than they had been in the 1970s (Fortune 1989). The second paper reviewed the recent literature on stock market efficiency, concluding that the Efficient Market Hypothesis, widely held in the 1970s and early 1980s, is not supported by the evidence (Fortune 1991). The existence of significant inefficiencies suggests that fundamentals do not play as central a role in market performance as has been thought.

The purpose of this article is to investigate the possible reasons for, and public policy responses to, very sharp short-term declines in stock prices. The focus will be the Crash of 1987, the most prominent stock market decline experienced in several decades. Of particular concern will be the role played by fundamentals and market mechanisms in this event, and the effects of recent financial innovations on the depth of the Crash.

This effort has not uncovered the "smoking gun" that would make the Crash a clearly understood phenomenon. In part, the inability to find "the" reasons for the Crash stems from the unique character of the experience; it does not allow easy generalizations. The Crash was the economic equivalent of a "hundred-year storm," a dramatic event on a scale beyond the capacity of established protective mechanisms, which occurs so rarely that its ultimate causes are often poorly understood.

A second reason for difficulty in understanding the Crash is that it was not a rational phenomenon, capable of being understood with the standard tools of economics. Some markets failed to perform properly, and these probably exacerbated—but did not create—a situation that turned into a panic. While public policy responses must be devoted to improving the functioning of those markets, the recurrence of a hundred-year storm cannot be avoided indefinitely.

The article begins with an introductory section that describes the Crash of 1987 and its history. Section I reviews daily volatility of stock prices in the 1980s. Section II discusses some possible fundamental causes of the Crash, including the filing of legislation to limit tax benefits of takeovers, a rise in interest rates, and the end of a speculative bubble. Section III discusses those features of the securities markets that have been blamed for the Crash, namely program trading, portfolio insurance, and index arbitrage. The next section discusses the efficacy of proposed policy responses to the Crash. The paper concludes with a brief summary.

The October 1987 Crash was the economic equivalent of a "hundred-year storm," an event beyond the capacity of established protective mechanisms.

The Chronology of the Crash

The peak before the crash occurred on Friday, October 2 for both the S&P 500 and the Dow Jones 30 Industrials. The following week saw the S&P 500 fall over 5 percent, and in the period from October 12 to 15 it fell an additional 9 percent. The latter four trading days were chaotic. The strong downward trend was accompanied by a high volume of trading in S&P index futures, and during brief periods stocks' futures prices were lower than cash prices. This "backwardation" provided a strong signal to sell stocks. By the Friday close the S&P 500 had fallen by 5 percent, and the S&P 500 index futures price was about equal to the S&P 500 index.

Sell orders accumulated over the weekend, and on Monday, October 19, opening prices were sharply lower. Almost 200 stocks failed to open on time because of order imbalances. Selling pressure built up as futures contracts continued to sell below stock prices. Long delays arose in the execution of sales, breaking the link between prices at the time orders were submitted and final execution prices. By the end of the day, the Dow Jones 30 index was down 508 points or 22.6 percent, with trading volume over three times that for a normal heavy day.

High anxiety about the market was widespread on Tuesday, October 20. Overnight the Nikkei 225 index had fallen over 13 percent, and by the New York open, the London FTSE was down sharply. Before the open, the Chairman of the Federal Reserve Board, Alan Greenspan, announced that the Fed would provide "a source of liquidity to support the economic and financial system."

The open on Tuesday saw a significant excess demand for stocks and within one hour the Dow Jones 30 rose by 200 points. While initially the S&P 500 index futures contract rose sharply, by 10:00 a.m. it began a fall that continued until noon, at which time trading in the S&P 100 and S&P 500 index futures contracts was halted by the Chicago Mercantile Exchange because trading had been halted in a significant number of S&P 500 stocks. At 1:00 p.m. the Chicago Mercantile Exchange restarted trading in stock index futures contracts and, during the afternoon, stock prices recovered and futures prices remained above the lows experienced at midday. By the end of the day, both the S&P 500 and the Dow Jones 30 were above their opening levels. Throughout the day the futures market remained at a significant discount to stock indices.

I. Short-Run Stock Price Movements in the 1980s

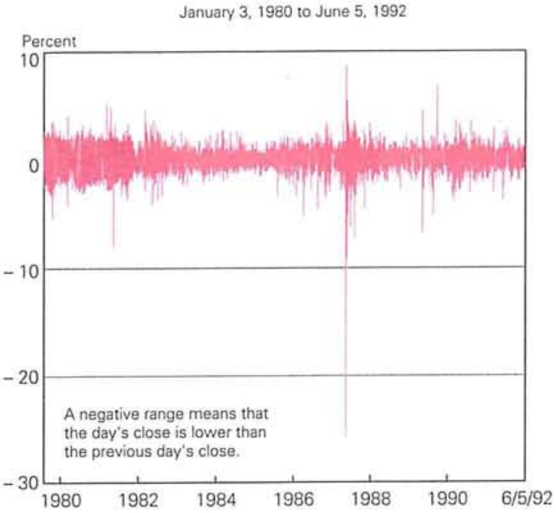
The adjusted intraday trading range of the Standard & Poor's 500 Composite Price Index provides a useful measure of very short-run price variability in the stock market. The adjusted intraday trading range is calculated by dividing the daily range (S&P 500 high less S&P 500 low) by the previous trading day's closing price. It can be interpreted as the difference between the daily "high" and "low" percentage changes in price. For example, an adjusted intraday range of 5 percent means that if the S&P 500 had been sold at its daily high, the percentage change over the previous close would have been 5 percent greater than if it had been sold at its daily low.

Overview of Daily Volatility: 1980 to 1992

Figure 1 shows the adjusted intraday trading range for the 3,141 trading days in the period January 3, 1980 to June 5, 1992. Though the range is, by definition, always a positive number, Figure 1 reports

Figure 1

Adjusted Intraday Price Range, S&P 500



Source: Author's calculations.

the trading range as negative on "down" days (when the close is below the previous close). This distinguishes between trading ranges on "up" and "down" days. No trend in daily volatility is apparent. The 1980–1981 period shows higher than normal volatility, but this is followed by abnormally low volatility during the bull market of 1982 to 1986. The most prominent revelation in Figure 1 is the occurrence of rare, unsustained bursts in daily volatility.

The analysis reported in Box I, "Time Series Analysis of Adjusted Intraday Trading Range," suggests several observations about the behavior of short-term volatility. First, no apparent trend in volatility occurred over the 1980s. Second, while days of high volatility tend to be followed by other high-volatility days, the effect disappears quickly. Third, volatility is particularly high on "down" days. Finally, volatility appears to be high after lengthy periods of closed markets.

The implication that trading halts will increase subsequent volatility must be accepted with caution, however. These halts are not the same as halts due

Box I: Time Series Analysis of Adjusted Intraday Trading Range

This box reports a simple ARIMA model of the *logarithm* of the absolute value of the adjusted intraday trading range.[1] An "intervention variable," named "DOWN" and defined as +1 on "down" days and zero on "up" days, was added to assess the possible asymmetry of the trading range on "up" and "down" days. In addition, two dummy variables were introduced to reflect breaks in market trading due to weekends or holidays: "BREAK1_2" is +1 when a trading day has been preceded by a one- or two-day break, such as weekends or holidays (0 otherwise), while "BREAK3" is +1 following a three-day break (0 otherwise). Experiments revealed that an ARMA(1,1) process captured the data. The results, with t-statistics in parentheses, are:

$$\text{(1)}\quad \log(AITR_t) = +.0274 + .9898 * \log(AITR_{t-1})$$
$$\qquad (.31) \quad (327.9)$$
$$+ \epsilon_t - .9075 * \epsilon_{t-1} + .0271 * BREAK1\_2_t$$
$$(-99.3) \qquad (1.16)$$
$$+ .1665 * BREAK3_t + .0659 * DOWN_t$$
$$(2.53) \qquad (3.53)$$
$$R^2 = .27 \qquad Q(24) = 21.47\ [.37]$$

The first part of this equation reports the Autoregressive–Moving Average estimates. These indicate significant autocorrelation in the adjusted intraday trading range: a surprise in the trading range has effects on future trading ranges. However, these effects dissipate rather quickly.

The coefficient on DOWN indicates that days of downward price movements (close-to-close) tend to have higher trading ranges than "up" days, confirming the notion that "crashes" tend to be accompanied by particularly high trading ranges. BREAK1_2 is positive but not statistically significant, while BREAK3 is both positive and statistically significant. Thus, it appears that long breaks in trading—at least of the prescheduled variety, like three-day weekends—are followed by a higher trading range.

[1]Because the adjusted intraday trading range (AITR) is necessarily non-negative, it cannot conform to the assumption of a normal distribution. The use of log(AITR) is a transformation that results in a variable more likely to conform to the normal distribution. As a result, the statistical properties of the ARIMA estimates for log(AITR) are more desirable than estimates for AITR.

directly to stock market performance (price crashes, execution backlogs, general chaos). Too few of the latter type of halt have occurred to allow generalizations. Even so, calendar and time-of-day halts are of some interest.

To investigate the behavior of stock markets during crashes, criteria to identify a crash must be chosen. In order to focus attention on the most prominent episodes, this study has identified all trading days between January 3, 1980 and June 5, 1992 that meet two criteria: (1) the day is a "down" day, in other words, the closing price is less than the previous close, and (2) the absolute adjusted intraday trading range was at or above the 95th percentile of all 3,141 values in the sample. Table 1 reports the 21 days that meet those two criteria. The range of 25.74 percent on October 19, 1987, is clearly the most extreme intraday volatility in the sample, dwarfing the next highest range (9.21 percent), which occurred one week later.

Table 1
Days of Stock Market Crash[a]
January 3, 1980 to June 5, 1992

Date	Adjusted Intraday Trading Range (AITR) (Percent)	Daily Price Change (Close to Close) (Percent)
19 October 1987	−25.74	−20.47
26 October 1987	−9.21	−8.28
21 October 1981	−8.15	−.15
08 January 1988	−7.44	−6.77
13 October 1989	−6.81	−6.12
30 November 1987	−6.32	−4.18
22 October 1987	−6.28	−3.92
16 October 1987	−6.15	−5.16
27 March 1980	−5.43	−.47
11 September 1986	−5.27	−4.81
26 January 1990	−5.26	−.09
03 November 1987	−5.17	−1.93
14 April 1988	−4.70	−4.35
23 January 1987	−4.65	−1.39
13 August 1980	−4.20	−.87
25 October 1982	−4.12	−3.97
03 December 1987	−3.86	−3.53
03 August 1990	−3.85	−1.88
15 November 1991	−3.80	−3.66
14 December 1982	−3.76	−1.83
08 March 1982	−3.75	−1.83

[a]A crash is defined as a "down" day for the S&P 500, with an AITR at or greater than the 95th percentile. For the 3,141 trading days in the sample, the median AITR was 1.10 percent, the interquartile range was 0.73 percent to 1.77 percent, and the 95th percentile was 3.72 percent.

The adjusted intraday trading range is the difference between the "high" daily percentage change (intraday high vs. previous close) and the "low" daily percentage change (intraday low vs. previous close). A positive (negative) sign indicates that the day's close was higher (lower) than the previous day's close.

II. Did Fundamentals Trigger the Crash?

One of the fundamental factors cited as a trigger was the *filing of takeover legislation* to limit tax benefits of corporate takeovers. This legislation, introduced on October 13 and approved by the House Ways and Means Committee on October 15, would have eliminated the deductibility of interest paid on debt issued for takeovers and other corporate restructurings.[2] What role did this play in the Crash?

The Brady Commission Report shows that the prices of stocks that were prime candidates for takeover fell sharply relative to the S&P 500 in the week before Black Monday. However, these takeover stocks had also outperformed the market for the year: between December 31, 1986 and mid October 1987, an index of eight takeover candidates had risen by over 70 percent, while the S&P 500 had risen by roughly 25 percent. Unfortunately, a decline of target firm prices relative to the market does not establish causation—an equally plausible hypothesis is that target firms are more volatile than the market and would suffer more in a down market, just as they did better in up markets.

A more subtle approach to assessing the role of takeover legislation is provided by the event study of Mitchell and Netter (1989), who identify five announcements associated with the anti-takeover legislation that would have affected takeover stock prices. Two that would have depressed prices occurred on October 13 and October 15, when the market first learned of the filing and subsequent approval of takeover legislation by the House Ways and Means Committee; these would affect trading on October 14 and 16. The other three took place after the crash, on October 29, October 30, and December 16. Public announcements on these dates concerned the moderation of the proposed legislation's restrictions on takeovers and the eventual loss of support for the legislation; these announcements would have led to increases in takeover stock prices.

Mitchell and Netter found that the rate of return on the S&P 500 Index on those five days conformed to

[2] The legislation limited interest deductibility for acquisition of a majority interest to $5 million per year. It also eliminated entirely any deductibility for hostile acquisitions of over 20 percent of a target's stock.

the predicted effects of the announcements on all five occasions: unusually negative returns on the first two events were followed by unusually positive returns on the next three events. They also found that target firm stock returns conformed to the predictions, with even stronger responses in returns. In addition, information on transactions by risk arbitragers indicates that they were responding to the information.[3]

It comes as no surprise that risk arbitragers respond to information on the tax benefits of takeovers, or that the prices of stocks of candidates for takeover are also sensitive to tax benefits. However, it is more difficult to understand why the general market, measured by the S&P 500, should be so sensitive to takeover legislation. Mitchell and Netter argue that the possibility of takeover is a way of dealing with agency problems: a reduction in the probability of a takeover allows management to ignore the interests of shareholders. Hence, shareholders of all firms benefit from low barriers to takeover. Even so, it seems unlikely that the general market could be so dominated by news that might affect only a select category of stocks.

A second candidate for triggering the Crash is *interest rate increases subsequent to a poor international trade report*. In August, when stocks were at an all-time high, the 30-year Treasury bond yield had averaged below 9 percent. But by Tuesday, October 13, Treasury bond yields had closed at 9.92 percent. By Wednesday's close—after the merchandise trade balance report in the morning—Treasury bonds were at 10.12 percent. As the decline unfolded, Treasury bond yields continued to increase until the close on Friday, October 16, at a rate of 10.24 percent.

Equation (1) is the most commonly used stock pricing model, for which m is the price-earnings multiple, r is the rate of discount, and g is the anticipated growth rate of earnings per share.

$$m = 1/(r - g) \quad (1)$$

According to this model, the proportional change in the multiple when the rate of discount changes is $dm/m = -m * dr$. The rise in the long-term Treasury bond yield from Tuesday to Wednesday was 0.20 percent (or 0.0020). At a multiple of 20 (the September 1987 average for the S&P 500), this implies a decline in stock prices of about 4 percent during Wednesday; the actual decline in the S&P 500 was about 3 percent. The model implies a stock price decline of about 6.4 percent from Tuesday through Friday; the actual decline was about 10 percent.

Thus, significant increases in long-term interest rates provide a plausible explanation of the trigger for the general stock price decline in the week prior to Black Monday. These increases were largely the result of an adverse trade balance report and the consequent loss of confidence in the dollar and in dollar-denominated securities.[4]

Yet another fundamental factor cited as a trigger for the crash is a *worldwide downward revision of expectations* that affected global stock markets. Roll (1988) argues that both the initiation of the U.S. stock market crash, and much of its depth, can be "ascribed to the normal response of each country's stock market to a worldwide market movement." In support of this, Roll notes a positive correlation among returns on common stocks in 22 countries from 1981 to September of 1987, with October of 1987 being the only month in that period when all 22 stock markets declined. Thus, he concludes, a general collapse in global expectations in a world of interconnected stock markets explains the October Crash.

The proximate cause of the Crash was the sharp increase in interest rates, combined with uncertainty about foreign holdings of U.S. securities, that followed the October 14 merchandise trade report.

Roll rejects institutional arrangements as a primary cause of the global crash, but he does examine the relationship between the magnitude of the crash in each country and the existence of several institutional arrangements. The results, while inconclusive, are interesting: countries with continuous auction markets tended to fare worse than countries with a specialist system, and countries with computer-directed order systems tended to fare better than those with manual systems. Thus, the United States

[3] Risk arbitrage is the term applied to purchase or sale of stocks in anticipation of mergers and acquisitions.

[4] Long-term interest rates had been increasing since January, while stock prices also increased. Over this longer period, it seems likely that earnings growth anticipations were the primary source of stock price increases. However, over the few days in October, earnings expectations were probably constant, so interest rates can be isolated as a factor.

should—and did—experience a smaller collapse in stock prices than other countries. Roll found no significant relationship between the magnitude of the crash and the existence of margin requirements, trading in stock index options or futures, or price limits.

A global speculative bubble might well have existed prior to October of 1987, but Roll's argument lacks one convincing detail: an indication of why global expectations should have been revised so sharply. Roll's hypothesis does serve, however, as a reminder that U.S. stock markets are connected with markets in other countries.

While fundamental factors may have played some role in triggering the October 1987 Crash, it seems clear that the magnitude of the Crash was far greater than fundamentals would indicate. This conclusion is supported by evidence on insider trading around the time of the Crash. Presumably, corporate insiders will be able to judge the fundamental values of their firm's shares, and will be net sellers of their firm's shares in the case of a downward revision in fundamental values. Seyhun (1990) examined monthly Securities and Exchange Commission (SEC) data on insider transactions for over 6,000 firms from 1975 to 1988, and found that during October of 1987 insider purchases were unusually high and insider sales were unusually low. This was true of top executives as well as lower-level management, and it was true of firms identified as takeover targets as well as other firms. Fundamental factors clearly were not recognized as a factor in the Crash by those best positioned to identify them.

Thus, we conclude that the magnitude of the decline in stock prices was considerably greater than fundamental factors can explain, and that an understanding of the extent of the Crash requires an examination of the non-fundamental factors that existed at that time.

Was the Crash the End of a Speculative Bubble?

One explanation of the Crash states that it was the inevitable consequence of unprecedented, and unwarranted, high stock prices. In short, the Crash occurred because an inexplicable boom had preceded it. This explanation is, of course, inconsistent with the notion of stock market efficiency.

The data are certainly consistent with this hypothesis. Prior to the Crash of October 1987, the stock market had been rising sharply: over 1986 the S&P 500 had risen by 14.6 percent, well above the normal rate of increase, while from January to the October 13 peak, the S&P 500 rose at an annual rate of 33.5 percent! At the time considerable debate took place about the reasons for this, but most financial economists considered the market to be unsustainably high.

This has led some observers to conclude that there was a *speculative bubble* in stock prices. The concept of a "rational" speculative bubble has been discussed elsewhere (Fortune 1991). This type of speculative bubble is an economic concept that should be distinguished from the layman's definition of a speculative bubble, which rests on hindsight: a lay interpretation of a bubble is merely what happens before a crash! A rational speculative bubble exists when asset prices become separated from fundamental values, *and* when investors believe that this will continue. Note that transitory departures from fundamental values, due, perhaps, to changes in conditions of liquidity, to adverse behavior of market-makers, or to "uninformed" traders' misperceptions about price, are not speculative bubbles. A necessary characteristic of a rational speculative bubble is that it be self-fulfilling, that either investors are not aware that it exists and so behave in ways that perpetuate it, or investors are indifferent to the existence of the bubble because they believe that a "greater fool" will rescue them from the consequences of overpayment.

The data are consistent with the hypothesis that the Crash occurred because an inexplicable boom had preceded it.

Rational bubbles, which can exist even in the presence of rational expectations about future dividends and earnings, have received a great deal of attention from theorists in recent years. While their existence is consistent with modern financial theory, it is extremely difficult to actually determine whether a bubble is present, that is, to distinguish between the part of the price that is due to fundamental value and the part due to the bubble.

The difficulty can be seen as follows. The standard discounted cash flow theory of stock prices results in the following description of the process generating movements in fundamental values: $P^*_{t+1} = (1 + r)P^*_t - E_tD_{t+1}$, where r is the required return on stocks, P^*_t is the fundamental value, and E_tD_{t+1} is the dividend expected next period (the expectation formed in this period). Now assume the existence of a bubble, denoted by B_t. Then the price of the stock will be $P_t = P_t^* + B_t$. If the bubble is self-fulfilling, investors must expect it to earn a normal rate of return. Hence, it would be expected to grow at the rate r, and $E_tB_{t+1} = (1 + r)B_t$. The stock price cum bubble will be $P_{t+1} = (1 + r)P_t - E_tD_{t+1}$, which is precisely the same path as the fundamental value. Thus, investors will never know that a rational bubble exists, because prices with a rational bubble follow exactly the same process as prices without a bubble. Furthermore, even if investors believe that a bubble is present, they will be willing to pay the higher price because they believe that the premium will earn the required rate of return.

Rational bubbles are subject to some restrictions. First, as noted above, they must increase at the required rate of return on stocks. Second, bubbles can never be negative. To see this, note that if a stock pays cash dividends the rate of increase of the stock price will normally be less than the rate of increase in the bubble. The bubble must then be an ever-increasing portion of stock price. If a bubble were negative, eventually the stock price would go to zero. No investor would knowingly hold a stock with a negative bubble because it will eventually become valueless; thus, any indication of a negative bubble must be immediately self-correcting.

Efforts to determine whether rational bubbles occur in stock prices have been inconclusive, though most studies have not supported bubbles. However, the focus has been primarily on the existence and continuation of rational bubbles, rather than on the bursting of bubbles. The bursting of a bubble is an infrequent occurrence, which economists are not well equipped to explain.

III. Market Factors: Program Trading, Portfolio Insurance, and Index Arbitrage

While fundamental factors can explain part of the price decline that occurred during the Crash, they cannot account for its most dramatic and alarming features, namely the panic selling and the precipitousness of the accompanying price fall. Program trading in general has been widely blamed in the aftermath of the Crash, though it is not clear whether the bad press was due to the poor popular reputation of futures-related trading, or to the notion of computer generation of orders. In fact, program trading is not the use of computers to initiate orders. The New York Stock Exchange (NYSE) defines a program trade as an order of $1,000,000 or more involving at least 15 stocks. While program trades are often initiated and transmitted via computers, and are usually associated with derivative securities such as stock index futures, no clear association exists between the volume of program trading and stock market volatility. A causal connection is even less clear. This has been established in a number of studies.

Fundamental factors cannot account for the panic selling and the precipitousness of the accompanying price fall during the Crash.

One prominent example is Grossman (1988). Using daily data for 1987, Grossman found that SuperDOT volume was positively correlated with market volatility. (SuperDOT is a NYSE computer order and transaction system.) However, he found no relationship between stock market volatility and program trading volume, using several measures of program trading. Thus, high-volume days tend to be days with high volatility, and these are also days with more intensive use of SuperDOT. But the volatility-program trading nexus appears to be absent.

Most program trading is done either for the purpose of limiting risk through "portfolio insurance" or for index arbitrage; both involve trading in stock index futures. While program trading in general does not appear to be at fault, both the Brady Commission (1989) and the Securities and Exchange Commission (1988), having carefully examined the chronology of the Crash, concluded that the problem lay in a destabilizing interplay between index arbitrage and portfolio insurance strategies involving index futures. This view has become known as the *Cascade Theory* of the Crash. While an analysis of various portfolio insurance strategies and their effectiveness is outside the scope of this article, Box II, "Stock Index Futures, Hedging, Portfolio Insurance, and Index Arbitrage" provides background to help the reader understand the mechanics of the Cascade Theory.

Box II: Stock Index Futures, Hedging, Portfolio Insurance, and Index Arbitrage

Stock Index Futures

A futures contract on a stock index is an agreement for the purchase or sale of that index at a specified future date and at a price determined at the time the contract was made. The first stock index contracts were approved for trading in 1982. At present, five stock index futures contracts are traded on several different exchanges: the S&P 500, traded at the Chicago Mercantile Exchange; the Major Market Index, traded at the Chicago Board of Trade; the NYSE Composite Index, traded at the New York Futures Exchange; the Value Line Index (Kansas City Board of Trade); and the Nikkei 225 Index, traded at the Chicago Mercantile Exchange. By far the most active trading is in the S&P 500 Stock Index futures contract, initiated in April of 1982. Open interest of 153,853 contracts on July 20, 1992 was six times the open interest on the next most popular stock index futures contract (the Nikkei 225, with open interest of 26,091).

For example, consider the Box Table, which shows the report of closing prices for the S&P 500 Index futures contract on Monday, July 20, 1992, when the index itself closed at 413.75. (The dollar value of a contract is 500 times the index.) If a trader bought a September S&P 500 Index futures contract at its closing price of 413.95 per unit, he would be obligated to take delivery of 500 units of the S&P 500 at the expiration date for a total cost of $206,975 (= 500 × 413.95). The profit or loss on that position will depend upon what happens to the index. If, for example, the index rises to 420, the investor can exercise the futures contract by taking delivery of 500 units of the S&P 500 at $206,975, then selling these units for $210,000 (= 500 × 420), realizing a net profit of $3,025. However, if the index falls to 410 he can take delivery of $206,975 of stocks that he can sell only at $205,000, a loss of $1,975.

Dynamic Hedging with Stock Index Futures

Dynamic hedging is the use of index futures, as well as other derivative instruments such as options, to hedge the risk of the stock portfolio. In practice, stock index futures are the least-cost method of risk reduction because they require essentially no investment. Suppose a financial institution wants to hedge the value of its stock portfolio by selling futures against the S&P 500. The first step is to calculate the number of the units of the index whose price variation can be offset by one futures contract, usually known as the hedge ratio, or delta (Δ). In order to fully hedge a portfolio of the S&P 500, the investor would sell $\Delta/500$ S&P 500 futures (each futures contract is for 500 units) for each unit of stock held.

The futures price is, in principle, the expected spot stock price at the expiration of the futures contract. The expected spot price, in turn, is the current spot price times an expected growth factor, which is the excess of the required return on the stock over the dividend yield (r − d), where r is the required return on the stock, and d is the dividend yield. If T is the time to expiration of the futures contract, the futures and spot prices are related by the valuation equation $F = Se^{(r-d)T}$. Hence, $\Delta = -(\delta F/\delta S) = -e^{(r-d)T}$ and insuring a unit of the stock index requires selling $[e^{(r-d)T}/500]$ S&P 500 futures contracts.

The ability to convert risky portfolios to riskless portfolios using derivative securities is not guaranteed. Dynamic hedging is designed to deal with normal "small" fluctuations in stock prices. If "jumps" in stock prices occur, the average delta will differ from the marginal delta derived above, and the method will fail to protect the portfolio from the price decline.

Portfolio Insurance

Portfolio insurance is a set of strategies de-

Stock Index Futures Trading, S&P 500 Index (CME)—$500 x Index
Monday, July 20, 1992

	Open	High	Low	Settle	Change	Open Interest
September 92	411.50	415.00	409.75	413.95	−1.10	148,496
December 92	411.30	415.30	410.10	414.35	−1.15	4,506
March 93	415.25	415.25	411.00	415.05	−1.10	751
June 93	—	—	—	415.85	−1.15	102

Estimated Volume 55,740; Friday Volume 51,754; Open Interest 153,853, +1,051 signed to prevent the value of a portfolio from falling below a prespecified floor at a specific point in time. For example, a financial institution with a portfolio currently worth $110 million might wish to ensure that its portfolio value is at least $100 million at year end. This can be done by periodic shifts of the portfolio between stocks and cash in response to actual stock prices. For example, if the stock market has risen, a larger portion of the portfolio can be invested in stock with confidence that the $100 million floor will be achieved. If, on the other hand, stock prices have fallen, the institution will invest a smaller portion in stocks and a larger portion in riskless securities, in order to protect the portfolio from falling below the $100 million floor.

Portfolio insurance's implications for market dynamics are a subject of considerable concern. Increases in stock prices lead the insuring institution to buy more stocks, while decreases in stock prices result in sales. Thus, the cyclical movements of the stock market are exacerbated by portfolio insurance. This dynamic portfolio reallocation is the source of the charge that portfolio insurance exacerbated the Crash in October of 1987.

Index Arbitrage: the Link between the Stock Index and Index Futures

Riskless index arbitrage occurs when a trader simultaneously buys (or sells) the individual shares in the S&P 500 in proportions indicated by market capitalization, and sells (or buys) an S&P 500 futures contract. The gain or loss from doing this is the difference between the futures price and the cash price; this is called the *spread* or *basis*.[5] For example, if the cash price of the S&P 500 shares is 420.55 and the S&P 500 index futures contract can be bought at 422.40, the spread is +1.85. In the absence of transactions costs, a spread of +1.85 means that a trader can, with certainty, make a profit (gross of transactions costs) of $925 (= 500 × 1.85) by buying 500 units of the S&P 500 and selling one S&P 500 futures contract. Thus, a positive spread provides an incentive to buy the S&P 500 in the spot market and sell the S&P 500 futures contract. A negative spread is an incentive to sell in the cash market (or sell short) and buy futures contracts.

In a world with no transactions costs or carrying costs, index arbitrage will ensure that the spread is zero. However, transactions costs prevent riskless arbitrage: one does not actually trade in all 500 stocks in the exact proportions needed to replicate the S&P 500 because of the commissions and other costs (such as bid-asked spreads) that must be paid. Instead, index arbitragers trade portfolios with a relatively small number of stocks that are highly correlated with the S&P 500. As a result, index arbitragers face *basis risk*, in the form of imperfect correlation between the S&P 500 and the portfolio they choose to trade. This basis risk must carry a reward, and the reward is in the form of a positive spread.

Carrying costs, such as the interest forgone on cash purchases net of dividends received, also induce a positive spread. As seen above, the futures and current stock prices are related by $F = Se^{(r-d)T}$, where r is the rate of interest, d is the dividend yield on the S&P 500, and T is the time to expiration of the contract. Hence, as an approximation, $[(F - S)/S] = (r - d)T$.[6] Because $r > d$, one should observe $F > S$, or a positive spread, even in market equilibrium. When the spread is positive by an amount equal to the cost-of-carry, the futures-spot relationship is in equilibrium and the markets are said to be "carry" markets, or to be in "contango." Carry, or contango, means that the trader experiences a net profit on the arbitrage equal to the costs of carrying the position. Clearly, if the spread is more positive than the cost-of-carry, index arbitragers will buy spot and sell futures until contango is created.

The atypical situation of a negative spread is called "backwardation." Backwardation is not an equilibrium situation because the index arbitrager has an incentive to sell (or to short) the stock index and buy futures. Thus, while a contango market *might* be in disequilibrium, a backwardation market *will* be in disequilibrium. The rational response to backwardation is to sell long positions in stocks, to sell stocks short, and to buy futures, thereby eliminating the disequilibrium. However, as we shall see, lengthy periods have occurred when the futures market was in backwardation with no apparent move to correction. This "mystery" plays a central role in understanding the Crash of 1987.

---

[5]The basis is typically defined as the cash price less the futures price, while the spread is the futures price less the cash price. However, it is common to use the terms spread and basis interchangeably and define both as futures price less cash price.

[6]Converting the futures valuation equation to natural logarithms gives $\ln(F/S) = (r - d)T$. But $\ln(F/S) = \ln\{1 + [(F - S)/S]\} \approx [(F - S)/S]$. Hence, as an approximation we can say that $[(F - S)/S] = (r - d)T$.

According to the Cascade Theory, the Crash began with a shift in fundamentals in the week before Black Monday but gained a momentum unrelated to any influence of fundamentals. The scenario goes something as follows:

- The initial decline of stock prices caused portfolio insurance programs to sell index futures in an attempt to limit losses on stock portfolios.
- This caused futures prices to fall so far that they traded at a discount from the spot prices, resulting in the backwardation of the index.
- The fall in futures prices fed back into spot stock prices, causing them to fall even further, and triggering further portfolio insurance sales of index futures. This encouraged index arbitragers to sell stocks and purchase index futures.
- This had the effect of reducing stock prices even further and feeding back to further futures price declines from portfolio insurance as well as from downward revisions of expectations about stock prices.

There is considerable reason to be skeptical of this mechanism. First, the observation stressed by the reports of both the Brady Commission (1989) and the Securities and Exchange Commission (1988), that portfolio insurers were selling futures and stocks, does not mean that they were driving futures prices down to unreasonably low levels. Indeed, as we shall see, this does not appear to have been true. Second, October 19 was a day of panic, and significant order imbalances occurred in both stock and futures markets because of expectations of further price declines, which led traders and investors to implement the time-tested method of portfolio protection: bailing out. Attributing the problem to futures-related trading might be a case of blaming the thermometer for the fever.

An additional reason for some skepticism is the empirical evidence. A central feature of the Cascade Theory is that futures prices fall "too much" because of portfolio insurance, pulling stock prices down via a dynamic process of index arbitrage and portfolio insurance. However, Santoni (1988) presents evidence rejecting this. Using minute-by-minute data for the S&P 500 Index and the December 1987 S&P 500 Index Futures contract, Santoni examines the lead or lag relationship between the spot and futures prices on Black Monday. He finds that changes in futures prices tended to lead changes in spot prices. While this result is consistent with the Cascade Theory, it is also expected in an efficient market when new information has its first impact in the futures market. Thus, this does not establish that markets were performing improperly.[7] Santoni also finds, however, that one feature of the Cascade Theory is not supported by the data: changes in spot prices do not generate subsequent changes in futures prices. Thus, changes in the spot market do not "cause" futures market adjustments. This, of course, is not consistent with the Cascade Theory.

Valid criticisms can be made of Santoni's argument and his conclusion. For example, during Black Monday very long lags occurred in the reporting of stock trades because of the unprecedented volume of trades. This raises the possibility that his data are corrupted: if the time stamp on stock trades is delayed, stock price changes will be reported as occurring later than the true time. Futures prices are reported promptly. Thus, the true sequence of leads and lags could be the opposite of that shown by the data. Such mistiming of trades did occur during the Crash. In the absence of direct evidence that it was sufficient to corrupt the data badly, however, Santoni's results remain valid.

Was the Futures Market Really in Backwardation?

Perhaps the most unusual feature of the Crash was the severe backwardation in the futures market. This situation was taken as evidence of a breakdown in relationships among security markets, with the implication being that the primary problem was in the futures market. However, the discounts in futures prices could have occurred because futures prices fell too much, because stock prices fell too little, or because a statistical illusion made it appear that a discount existed when it did not.

The Brady Commission attributed the discounts to an excessive selling pressure in the stock index futures markets arising from portfolio insurance. To the extent that this is true, the opportunity for index arbitrage should bring the cash market down as well, transmitting excessive price declines in futures to excessive price declines in stocks. Since the Brady report, the futures markets have commonly been thought to have failed during the Crash.

An alternative hypothesis to explain the magnitude of the Crash suggests that the discounts on futures were not "real," but were a statistical illusion

[7] In an efficient market, in which new information is rapidly reflected in market prices, one would expect that futures prices would adjust more rapidly because of the lower costs of transacting in futures contracts, and because the spot index tends to adjust more slowly as a result of "stale prices."

resulting from "stale prices" arising from "nonsynchronous trading" of the stocks comprising the S&P 500 Index. In short, the discounts were smaller than they appeared, perhaps even nonexistent. This, it is argued, had two effects. First, the apparent backwardation incorrectly signaled that cash prices were going to fall even further, thereby inducing institutions and traders to sell stocks to avoid larger losses. Second, the backwardation provided an incentive for index arbitragers to sell stocks or sell short, thereby adding to the pressures on the stock market. These two effects, both of which would induce larger sell orders than appropriate, assume that traders were not able to correctly evaluate the true discount.

The phenomenon of nonsynchronous trading is an inevitable consequence of the way stock indices are computed. Not all stocks in, say, the S&P 500 are trading at each moment, so the index is computed using the price of each stock at its latest trade. Normally this creates no problems, though it does serve to explain why evidence shows that the S&P 500 is serially correlated at high frequencies (say, five-minute data) while the underlying stocks do not exhibit serial correlation.[8] However, under crash conditions the problem of stale prices can be severe. It is common for a significant number of stocks either to open late or to experience trading halts. For these stocks, the last price used is higher than the "true" current price, and the price decline is not reflected in the computed index until the stock begins trading. Thus, trading halts create lags in the decline of the stock index.

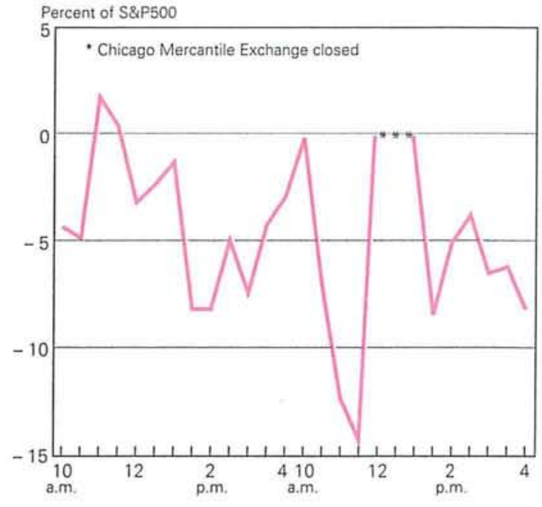
Table 2 shows the reported S&P 500 and the S&P 500 Index Futures prices at half-hour intervals on October 19 and 20, the days when the spread was negative. Figure 2 shows the futures-cash spread for these two days. This spread indicates that index arbitragers who relied on the *reported* S&P 500 index had a very strong incentive to close out their long positions in stocks and to sell short, while buying futures contracts.

Table 2 also shows the S&P 500 corrected for stale prices in two ways. The first correction (column 3) assumes that the price of a stock that is not traded is equal to the last trade price plus an adjustment equal to the proportional change in prices of traded stocks since the last trade of the untraded stock; this correction assumes that during nontrading periods, a stock's price followed the prices of traded stocks. The second correction (column 4) assumes that the price of an untraded stock is equal to the price at which it ultimately opened (if there was a delay in opening) or the price at which it reopened (if trading was halted after the open).

Figure 2

Discount on S&P 500 Futures Contract

October 19 and 20, 1987



Source: Securities and Exchange Commission (1988).

On October 19, some 57 S&P 500 stocks had delayed openings. But by noon, all but six of the S&P 500 stocks had opened, and any significant effect of stale prices due to trading halts had disappeared. Thus, by noon the "corrected" S&P 500 was very close to the reported index. In spite of this, the afternoon hours were all marked by backwardation in the futures market, providing a strong incentive to sell stocks short and to buy futures. Incorrect signals to index arbitragers from nonsynchronous trading do not appear to have been a major source of the backwardation: the backwardation appears to have been "real," in the sense that the reported S&P 500 index was accurate for most of the time.

October 20 was even more difficult as far as delays in opening and intraday closings were concerned. Indeed, the hours of 11:30 a.m. to 1:00 p.m. had more stocks not trading than at the open.[9] While

[8] Using last trade prices for infrequently traded stocks will introduce serial correlation into the index even if "true" prices were not serially correlated.

[9] Because of the significant nontrading during this period, the Chicago Mercantile Exchange suspended futures contract trading.

Table 2
Effect of Halt-Related Stale Prices on S&P 500 Index
Half-Hour Intervals October 19 and 20

Time	S&P 500 Non-Trades (1)	Reported S&P 500 (2)	Corrected S&P 500[a] Equal Proportion (3)	Corrected S&P 500[a] Open/Reopen Price (4)	December S&P 500 Index Futures (5)
Monday Oct. 19					
10:00	95	273.17	267.64	259.88	261.5
10:30*	73	265.77	262.11	254.21	253.0
11:00	37	258.38	257.34	254.41	263.0
11:30	12	263.85	263.79	263.33	265.5
12:00*	6	265.28	265.14	264.93	257.0
12:30*	2	259.89	259.71	259.71	254.5
1:00*	2	257.17	257.01	256.99	254.0
1:30*	0	255.70	255.70	255.70	235.0
2:00*	0	247.00	247.00	247.00	227.0
2:30*	0	245.00	245.00	245.00	233.0
3:00*	0	243.93	243.93	243.93	226.0
3:30*	1	235.78	235.74	235.77	226.0
4:00*	2	225.41	225.25	225.45	219.0
Tuesday Oct. 20					
10:00*	52	238.26	244.31	247.01	238.0
10:30*	19	245.16	245.71	245.35	228.0
11:00*	15	238.14	237.45	237.33	209.0
11:30*	38	223.78	222.75	222.21	192.0
12:00	63	221.39	221.97	219.84	closed
12:30	77	216.64	216.53	215.09	closed
1:00	57	228.39	231.06	227.64	closed
1:30*	33	225.87	225.97	225.88	207.0
2:00*	23	225.22	225.44	225.18	214.0
2:30*	17	227.95	228.32	227.89	219.5
3:00*	10	236.13	236.28	236.06	221.0
3:30*	3	240.20	240.21	240.18	225.5
4:00*	1	237.74	237.74	237.74	218.5

[a]The Corrected S&P 500 is done two ways: (3) Equal Proportion assumes that stocks not trading open with change from last trade equal to the proportional change in the traded stocks; (4) Open/Reopen assumes that stocks not trading have prices equal to the price at open/reopen.
*Times marked with an asterisk are times of backwardation (futures selling at discount from cash) using both measures of corrected S&P 500.
Source: Securities and Exchange Commission (1988), Chart 2–1, pp. 2–44, 2–45.

stale prices played very little role at any half-hour, the futures market was in backwardation at every half-hour during the day.

The above analysis, reported in SEC (1988), is based on rather crude methods. They are supported, however, by Harris (1989), who uses higher-frequency data (five-minute intervals) and a more sophisticated method of assigning prices to untraded stocks. Thus, the problem of *halt-related* nonsynchronous trading was severe around times of daily opening, but it disappeared fairly rapidly as delays in trading ended. The negative spread that was prominent during October 19 to 20 cannot be explained by stale prices caused by nonsynchronous trading.

While nontrading does not appear to have been an important reason for stale prices, a second reason for stale prices appears to have more power. This has to do with the existence of limit orders (see Box III, "Trading Terms") and with the potential for long delays in order submission when stock market volume is unusually high. If a specialist is flooded with sell orders, he will typically match those with limit buy orders in his Book. These limit buy orders were submitted at a time before the flood of sell orders, hence they do not reflect the sudden appearance of extreme pessimism. Thus, the reported price of the stock will remain high even though the "true" price is much lower. In effect, those who placed limit buy

Box III: Trading Terms

A *limit order* is a type of restricted order that sets a price limit which must be achieved. An order to "buy 500 BSX, limit 18," must be executed at a price of 18 or lower, while a limit sell order must be executed at the stated price or higher. If the broker who receives a limit order cannot execute it immediately with a floor broker or with the specialist in the stock, the order is placed on the specialist's Display Book, to be executed on a first come–first served basis when the limit can be met.

A *specialist* is a member of the New York Stock Exchange who buys and sells a specific stock for his own account in an attempt to maintain a fair and orderly market. The specialist also acts as a broker by bringing buyers and sellers together. In either capacity, the specialist provides quotes to the *commission brokers*, who take orders from their firm's trading desk or from registered representatives, and to the *floor brokers*, who can engage in transactions for their own benefit. These quotes, in the form of "BSX bid 17¼—asked 17¾," can be chosen from the Limit Book or, if the rules allow, the specialist can quote for additions to or sales from his own inventory.[10] The broker function of the specialist is to keep the book of limit orders from which quotes can be drawn. Thus, if the book has a high bid of 17 for BSX, and a low ask of 17¾, the specialist can quote that bid and ask; the specialist is then acting as an agent rather than as a principal. An additional important function of the specialist is to provide opening quotes at the beginning of each trading day. These can be difficult to construct when the book is thin or order imbalances have developed over the weekend or overnight.

Orders come to the specialist through two routes. First, and most common, is the commission broker, who approaches the specialist post for the stock and asks for quotes, but who might make a trade with a floor broker if that is more beneficial to the client. If these orders cannot be executed because of stops or limits, they are left on the specialist's book.

A second route, used primarily by office members on behalf of large accounts such as pension funds and mutual funds, is submission of orders through the *SuperDOT* system. (DOT refers to Designated Order Turnaround.) SuperDOT is a computer order and transaction system that has several components. The OARS (Opening Automated Report Service) component of SuperDOT accepts pre-opening market orders of up to 30,999 shares, which are electronically transmitted to the specialist for use in establishing opening prices. SuperDOT also accepts post-opening market orders of up to 30,999 shares and limit orders of up to 99,999 shares. These orders are electronically transmitted to the specialist's book or, in cases where the specialist does not have a Display Book, are printed on cards by high-speed printers on the floor. While these large orders can be carried manually to the specialist, SuperDOT normally provides a more rapid execution. SuperDOT originated in the early 1980s in response to the increasing institutionalization of trading, as mutual funds and pension funds became the primary traders and required a mechanism that could handle large orders quickly. In 1991, orders placed through SuperDOT (in number of shares) amounted to approximately 25 percent of NYSE volume.[11]

[10]BSX is the symbol for Boston Scientific Corporation, a medical devices firm recently listed on the New York Stock Exchange.

[11]SuperDOT trading can be compared to NYSE trading in a variety of ways. This study has chosen to use total orders (number of shares) placed through SuperDOT (both buy and sell orders) relative to total orders (number of shares) executed on the NYSE (both buy and sell). The latter is twice the reported NYSE volume. This corrects for the double counting in NYSE reports of SuperDOT trading.

orders are overpaying for the stock because they did not know that the sell orders would create an imbalance which would have allowed them to buy "at the market" at a lower price.

The effect of limit orders is, therefore, to create stale prices even though the stock continues trading. The illusion of a high stock index in the face of massive sell orders will be greatest when a crash is under way. Furthermore, the problem can be long-lasting if significant delays occur between the time a customer first initiates a limit order and the time it is recorded on the specialist's book. If long delays make the reported index very stale, traders will think the market is higher than it really is, and new limit buy orders that are placed will have too high a price limit. Thus, a continuing fresh supply of outdated limit orders can be generated, exacerbating the staleness in the reported index. This source of stale prices apparently was quite significant on October 19 and 20. The floor printers, which print execution order cards for the specialists, had a backlog of as long as 75 minutes on Black Monday, and electronic orders transmitted to the specialists' Display Books also were subject to significant delays. During the day the New York Stock Exchange requested that orders not be submitted through SuperDOT because it was so backed up, but the manual method involved even longer delays from the time a customer originated an order to the time it was executed.

In addition, the reports of executed trades were delayed because the cards describing them could not be filled out quickly enough. As a result, individual stock price results were delayed and investors had late information on them. Furthermore, traders did not know whether their limit orders had been executed, making it difficult to know whether they should be canceled or modified.

Kleidon and Whaley (1992) have demonstrated that stale limit-order prices were a significant problem on October 19 and 20. Five-minute price changes of individual stocks were not serially correlated during the Crash, but the S&P 100 and S&P 500 stock indices were serially correlated, a symptom of stale prices. While mild serial correlation in the index is normal, the extent of serial correlation was much greater on October 19 than during earlier trading days in October. The result was that "true" stock prices fell sharply, with considerable intraday volatility, while the reported stock price indices showed unusually smooth behavior on their downward trend.

The effect of limit orders is to create stale prices even though the stock continues trading, and this source of stale prices was quite significant on October 19 and 20.

It appears that informed traders were not fooled by the stale price problem. Kleidon and Whaley computed the values of the S&P 100 implied by the November 1987 S&P 100 stock index option contract. The results of their calculations for October 19, done at five-minute intervals, are reproduced here as Figure 3. While the implied S&P 100 tracked the reported S&P 100 quite well in the days prior to the Crash, on October 19 and 20 the implied index levels were far below the reported index levels. Thus, options traders appear to have been aware that the market was "really" trading at levels well below the reported levels. Because the implied S&P 100 index level on October 19 corresponded well with the S&P 500 futures price, it can be concluded that futures traders were not entirely fooled by stale prices.

Thus, significant discounts from the reported S&P 100 and S&P 500 indices appeared in both the options markets and the futures markets, which appear to have given more accurate estimates of the stock index than did the reported index.

Were Stock Index Futures Oversold?

The Brady Commission concluded that the futures markets failed to perform properly during the Crash and selected the stock index futures market as a significant source of destabilization. The futures and options markets appear to have reflected accurately the state of the stock market during the Crash, however, and the primary locus of market failure appears to have been in the New York Stock Exchange, where long delays in limit order submissions resulted in an illusion of discounts on futures contracts. This, in turn, gave incorrect signals to traders that the market was poised for further sharp declines. Option prices indicate that informed traders were not fooled, but the unprecedented high discounts on futures undoubtedly led the less informed traders, also called "noise traders," to engage in protective strategies such as outright sale of stocks.

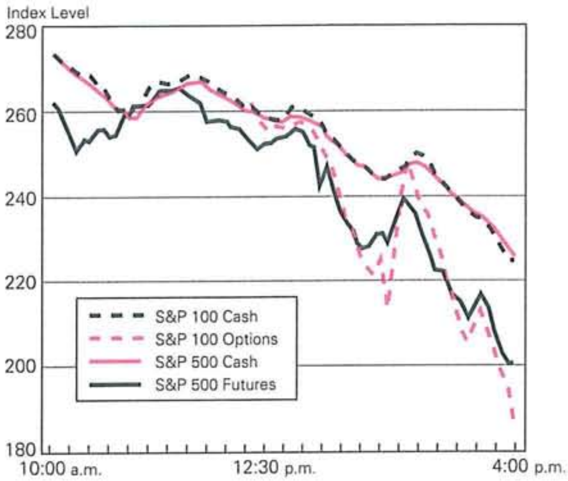
While the primary market failures appeared to be in the cash market, it is useful to ask how well the futures market performed its primary task of price discovery. Were futures prices during the Crash unreasonably low, in the light of actual market performance after the Crash?

If the futures markets were successful in serving their price discovery role, the October 19 and 20 futures price for the December 1987 contract would provide an optimal estimate of the actual S&P 500 on the expiration date, December 18. On the other hand, if futures markets were oversold, the futures price would provide an unusually low forecast of the

Figure 3

Actual and Implied Stock Index Levels

October 19,1987



Note: Index levels at five-minute intervals during the trading day, October 19, 1987. The figure contains the levels of the S&P 500 cash index and the December 1987 S&P 500 futures contract. The S&P 100 cash index level is normalized to the S&P 500 cash index level at 10:00 a.m. (EST). The implied S&P 100 index level is computed on the basis of November 1987 S&P 100 index option price quotes during each five-minute interval, and is normalized using the same proportionate adjustment as is used for the S&P 100 cash index.

Source: Kleidon and Whaley 1992.

December 18 index. Thus, a crude test of the oversold hypothesis is a comparison of the S&P 500 stock index futures price during the Crash with the S&P 500 stock index level on December 18, 1987, when the December S&P 500 futures contract expired.

On Friday, December 18, the S&P 500 closed at 249.16 after a day of trading in the 242.98 to 249.18 range. On October 19, the December S&P 500 futures price closed at 201.50, after trading in a range of 198.00 to 269.00. Thus, the October 19 closing price for the December S&P 500 index future was 47.66 points or about 19.1 percent below the actual December 18 index. This shortfall was certainly in the right direction, and of a magnitude to support the oversold hypothesis. But was it an unusual shortfall?

The analysis in Box IV, "Were Index Futures Oversold on Black Monday?" suggests that an underprediction of 47.66 points would occur about once every three years. While rare, it is not so rare as to suggest that the futures market was drastically out of line.

So What Really Happened?

While offering no definitive test of the hypothesis, this writer believes that the primary factor initiating the October 1987 Crash was the dramatic surge in stock prices that had occurred in 1986 and 1987. The resulting bubble set the stage for a sharp decline, and the proximate cause of the decline was the sharp increase in interest rates, combined with uncertainty about foreign holdings of U.S. securities, that followed the merchandise trade report on Wednesday, October 14.

A sharp increase in interest rates along with unsettling economic news, ending with an adjustment in the level of stock prices, is not a rare event, however, and cannot explain the depth and rapidity of price declines on October 19. What else was happening?

Although the major culprits are widely believed to be program trading, presumably encouraging floods of sell orders, and the existence of a variety of destabilizing strategies involving stock index futures, no evidence suggests that program trading was a causal factor in the Crash, nor did futures markets fail to perform their proper functions. The fundamental problems were largely in the cash market for stocks. The inability to expedite the large volume of orders led to a classic portfolio insurance strategy—if prices are falling and you do not know what is happening, get out! Thus, a large volume of sell orders begat a larger volume of sell orders, and longer backlogs of unfilled orders.

Combined with this was an important information problem. The backlog of unfilled limit orders resulted in stale prices, which made the reported index levels an unreliable measure of the state of the market. This led to an apparent backwardation in the futures market. Evidence from the options and futures markets suggests that many traders were aware of this problem: the implicit S&P index levels embedded in options prices were well below the reported index, and index arbitrage transactions were far smaller than the discount on futures would warrant.[12] But the discounts on futures contracts stood as a strong signal to uninformed traders that prices were headed lower still, hence encouraging the noise trading that created that very result.

[12] The relatively light volume of index arbitrage program trades was also due, in part, to NYSE admonitions against use of program trades during the Crash.

Box IV: Were Index Futures Oversold on Black Monday?

On October 19, 1987, the closing price for the December S&P 500 stock index futures contract was 201.50. The actual closing price for the S&P 500 on December 18, the expiration date, was 249.16, which is 47.66 points above the forecast of the index futures price on Black Monday. Does the forecast error of 47.66 points prove that index futures were oversold on Black Monday?

To answer this a simple statistical test can be employed. Let ${}_tF_{t+k}$ be the futures price at time t for a contract expiring at time t+k. Also, define S_{t+k} as the actual index level on the expiration date and ${}_tS_{t+k}$ as the rationally expected time t+k index level, with expectations formed at time t. Finally let $F = \ln(S_{t+k}/{}_tF_{t+k})$ be the measure of the forecast error.

The evolution of stock prices assures that $\ln S_{t+k} = \ln S_t + \Sigma\epsilon_{t+i}$, where each ϵ_{t+i} is the revision in the logarithm of the rationally expected price due to the arrival of new information at that time. Thus, one can derive the relationship $F = b + \Sigma_{i=1}^{k} \varepsilon_{t+1}$, where $b = [\ln({}_tS_{t+k}) - {}_tF_{t+k}]$ is the bias in the futures price forecast.

If the ϵ_{t+i} are independently and normally distributed with zero mean and standard deviation σ, then (F − b) is normally distributed with standard deviation $\sigma\sqrt{k}$. If σ were known, the statistic $(F - b)/\sigma\sqrt{k}$ would be a standard normal random variable. Because σ is not known, the sample standard deviation, s, must be used and the test statistic $(F - b)/s\sqrt{k}$ will be distributed as Student's t with the number of degrees of freedom for s.

In order to obtain an estimate of s, daily data for 2Jan1987 through 10Oct1987 were used to estimate a regression of the form $\Delta\ln S_t = \alpha + v_t$. Assuming that the logarithm of stock price at each time is the rationally expected value of the realization, the values of v_t are revisions in rationally expected values and can be used as a measure of the forecast revisions, ϵ. Thus, the standard error of estimate from this regression can be used as a measure of s.

There were 282 days in the sample, but only 196 trading days. The estimation procedure, which adjusted for holidays and weekends to estimate a daily standard deviation, generated a daily standard error of estimate of 0.0092. The gap between Black Monday and contract expiration on December 18 was 60 days, so the denominator in the test statistic is $s\sqrt{k} = (0.0092)(\sqrt{60}) = 0.0713$. Assuming that the futures price is an unbiased forecast (so b = 0), the numerator in the test statistics is $F = \ln(249.16/201.50) = 0.2123$. Thus, the value of the statistic is +2.98: a 47.66 point underprediction of the S&P 500 index is equivalent to a forecast error 2.98 standard deviations above the mean.

Because of the large sample size, the normal approximation to Student's t can be used. The probability of a standard normal variable of +2.98 or more is 0.0014, or 14/100 of 1 percent. With about 250 trading days per year, an underprediction of this size or larger would occur about once every three years.

IV. Policy Responses to the October 1987 Crash

The Crash resulted in a flurry of recommendations from both official and unofficial sources, each designed to limit the possibility of such a serious event recurring. At the time there was little understanding of the fundamental causes of the Crash. Hence, these recommendations were made in a near vacuum. The genes of the Crash still have not been isolated, so the proposals that have been adopted are not genetically engineered for a crash setting. The primary proposals have been of three types: use of trading halts, introduction of "circuit breakers," and introduction of margin requirements on derivative securities.

Trading Halts

Trading halts can occur at the discretion of the Exchanges or as the result of established rules. In the latter case, the halt is the result of circuit breakers, which will be discussed below. Discretionary trading halts are the subject of this section.

Trading halts have some clearly adverse consequences. First, because an important function of markets is to provide liquidity—the ability to execute transactions rapidly at appropriate prices—trading halts interrupt the normal functioning of markets. Thus, investors will pay lower prices for securities (require higher yields) if they believe that their ability to sell can be weakened by trading halts.

A second consequence of trading halts is that they can become self-fulfilling: if investors anticipate a trading halt, they will take evasive actions that trigger the halt. Thus, the prospect of a halt can create the certainty of the halt. This is particularly true of halts resulting from established rules (circuit breakers).

A third consequence of trading halts is the transmission of pressures to other markets as investors find substitute methods of achieving their goals. For example, a trading halt in stock index futures, as investors attempt to hedge their long positions, can induce larger sales of stocks in the cash market, driving the stock index down further. For example, on October 20, the period with the highest number of stocks not trading on the New York Stock Exchange was also the period during which the Chicago Mercantile Exchange halted trading in stock index futures. While it is widely reported that the CME closing was due to the number of halts on the NYSE, the direction of causation is not clear. In the same way, a halt in trading on the registered stock exchanges can transmit the excess sell orders to the futures market and to other cash markets, such as the over-the-counter market.

The case for halts is based on the concept of the "fog of battle." During periods of sharp price changes and, typically, a high volume of transactions, the information coming to traders is of low quality. They observe major price changes but do not understand whether they are permanent or temporary, due to fundamentals or to market overreaction. The result, it is argued, is that traders look to the recent performance of prices to form judgments about near-term performance: a decline in prices is extrapolated to continue into the future. As a result, markets become chaotic and price declines breed further declines. Furthermore, a natural response to confusion is to seek safety in riskless securities, thereby adding to sales of long positions and to the purchase of safe securities. Indeed, this seemed to characterize Black Monday, for while stock prices plunged, U.S. Treasury bond prices soared.[13]

In order to investigate the effect of trading halts on stock price volatility, it is useful to know whether the implied volatility during trading halts exceeds the normal volatility. For this purpose, a simple norm can be established: if σ is the standard deviation of, say, hourly prices during trading hours, then the standard deviation of prices over T hours of nontrading should be $\sigma\sqrt{T}$. If trading does not affect the volatility of stock prices, the volatility over a T-hour trading halt (from close to open) should also be $\sigma\sqrt{T}$. Higher volatility over halts indicates that trading reduces volatility; lower volatility over halts indicates that trading increases volatility.

The evidence on the effect of trading halts is, unfortunately, based on calendar or time-of-day events that are known in advance, such as holidays or weekends. In a well-known paper, French and Roll (1986) compared variability over weekends, midweek holidays, and holiday weekends with variability during trading sessions. The sample was all NYSE and AMEX stocks during the period 1963 to 1982. Defining price changes during trading as open-to-close, and overnight changes as close-to-open, they concluded that volatility over these calendar halts was considerably lower than volatility during trading sessions. Indeed, the differences were dramatic, with holiday weekends and normal weekends having about 10 percent of the normal volatility, while midweek holidays showed 27 percent of normal volatility. Thus, the act of trading itself appears to increase volatility in prices, suggesting that even under normal circumstances trading halts can be a stabilizing influence.

Trading halts interrupt the normal functioning of markets, and can become self-fulfilling.

On the other side, Amihud and Mendelson (1987) found that open-to-open price variation is significantly greater than close-to-close variation for the 30 stocks in the Dow Jones Industrials. In a second paper, Amihud and Mendelson (1991) examined the Tokyo Stock Exchange, which has two separate sessions in each day, hence a midday trading halt, and found the same result. These results suggest that the task of finding an opening price introduces variability, and that trading halts might increase volatility.

[13] Thirty-year U.S. Treasury bond yields were 10.25 percent at the Black Monday close but had fallen to 9.11 percent by Friday's close.

The results of French and Roll seem more relevant, because they deal directly with the implied volatility during halts. Neither approach, however, really gets at the main question of the effect of halts under chaotic trading conditions. At this point, no conclusion can be reached on the consequences of trading halts.

During the Crash several interventions occurred that might be similar to trading halts. The admonition not to use SuperDOT on October 19, and the CME closing on October 20, are examples. It appears likely that these were the wrong steps, and that they interfered with investors' access to timely information and trades in a fashion that increased the fog of battle.

Circuit Breakers

Circuit breakers consist of rules to halt trading or to alter the order process in a fashion designed to allow gathering of information. Among these rules are price limits, which prevent trading at prices sufficiently above or below the previous close, and trading halts, which prevent trading at any price.

Box V, "Circuit Breakers in Cash and Futures Markets," describes the circuit breakers adopted by the Chicago Mercantile Exchange and the New York Stock Exchange on October 20, 1988 and amended in 1990. The CME circuit breakers for the S&P 500 index futures contracts are based on price limits. First is a five-point open price limit on changes in the S&P 500 opening price over the previous close. This triggers a 10-minute delay in trading, designed to prevent chaotic openings that might result in inappropriate transactions in the cash markets as well as in the futures markets.

In addition, the CME adopted two other levels of price limits. Under the initial daily limit, if the S&P 500 futures price falls more than 12 points from the previous close, the 12-point floor must be maintained for 30 minutes or until 2:30 p.m. Chicago time. Also, a maximum daily price limit prohibits trading at futures prices more than 20 points below the previous close. Finally, if the NYSE declares a halt in trading, the CME also halts futures transactions until 50 percent of S&P 500 stocks have resumed trading.

The NYSE circuit breakers are more moderate. First, under Rule 80A, a 12-point decline in the S&P 500 futures price (which triggers a 30-minute floor on the CME) triggers two mild circuit breakers on the NYSE: (1) a "sidecar," in which program trades submitted to SuperDOT are delayed for five minutes before execution (manual program trades are not delayed), and (2) a prohibition on stop or stop-limit orders for a member firm's account for the remainder of the day. Also under Rule 80A, a 50-point change in the DJ 30 subjects index arbitrage orders to a "tick test": if the DJ 30 is down 50, sell orders must be executed on an up tick; if the DJ 30 is up 50 points, buy orders must be executed on a down tick.

Under rule 80B, there is a sequence of trading halts based on declines in the DJ 30. A 250-point decline triggers a one-hour trading halt, and a 400-point decline triggers a two-hour halt; these have never come into effect.

Circuit breakers are subject to the criticisms lodged against any trading halts. In addition, because they are triggered by clearly announced rules, they are more easily anticipated and potentially more likely to be triggered because of evasive actions. Several additional criticisms have been made. First, in order to work well they should be coordinated across markets: a circuit breaker tripped in one market should not allow trading to be diverted to substitute markets. The circuit breakers adopted in October of 1988 did not have that feature. This lack of coordination has been somewhat reduced by the amendments adopted in 1990. At present, a trading halt on the NYSE automatically creates a halt in trading on all other equity, index options, and index futures markets. However, the reverse is not true: trading halts in derivative securities are not necessarily matched by equity market halts.

Lack of coordination can exacerbate short-term volatility as the natural mechanism for inducing price stability—allowing competitive trading in substitute products—is eliminated. For example, if a sharp fall in the S&P 500 index futures contract initiates a futures trading halt, it also eliminates the ability to hedge a long position using futures, inducing investors to sell more sharply in the cash market.

A second problem with circuit breakers is that the triggers must be adjusted continuously as markets change over time. For example, in October of 1988 the five-point limit on open prices in the S&P 500 futures contract translated to a 1.8 percent change in futures price at an S&P 500 index level of 275; at the present index level of 415, a five-point limit translates to only 1.2 percent. Furthermore, a trigger of 250 points in the DJ 30 was equivalent to an 11.6 percent decline in October of 1988, but is only about 7.5 percent at present.

The need to adjust the trigger points is also affected by changes in relationships across markets.

Box V: Circuit Breakers in Cash and Futures Markets
(First Adopted October 1988, Amended December 1990)

New York Stock Exchange Circuit Breakers

NYSE RULE 80A

1. Trigger: *S&P Index Futures price falls 12 points* below previous close.
   Results: (i) *Sidecar*. Program trading orders submitted to SuperDOT will be routed to a separate file (the "sidecar") and held five minutes; there is no sidecar for manually transmitted orders. At the end of the five-minute period, the orders will be transmitted to the appropriate specialists.
   (ii) *Stop Order Prohibition*. New stop and stop-limit orders for a member firm's account are prohibited for the remainder of the day; stop and stop-limit orders for 2,099 shares or less submitted on behalf of an investor are allowed.

2. Trigger: *DJ30 moves ±50 points* from previous close.
   Results: *Tick Test*. Index arbitrage orders for component stocks in the S&P 500 can be executed only if they meet a "tick test"; this prohibits selling on a downtick or buying on an uptick. This tick test remains in effect for the remainder of the day or until the DJ30 moves back to ±25 points from the previous close. The tick test does not apply to index arbitrage orders submitted on exercise dates for liquidation of positions involving derivative securities.

NYSE RULE 80B

3. Trigger: *DJ30 index falls 250 points* below previous close.
   Results: *One-hour trading halt* on NYSE and all other equity, options, and futures markets. If the trigger was pulled between 3:00 p.m. and 3:30 p.m., the NYSE has the discretion to permit trading to reopen before 4:00 p.m.

4. Trigger: *DJ30 index falls 400 points* below previous close.
   Results: *Two-hour trading halt* on NYSE and all other equity, options, and futures markets. If the trigger was pulled between 2:00 p.m. and 3:00 p.m., the NYSE has the discretion to permit trading to reopen before 4:00 p.m.

Chicago Mercantile Exchange: Circuit Breakers for S&P 500 Index Futures Contracts

1. Trigger: *Open Limit*. S&P 500 futures opens ±5 points from its previous close.
   Results: 10-minute delay in futures trading.

2. Trigger: *Initial Daily Limit*. The S&P 500 Index Futures price falls 12 points below its previous close.
   Results: *Trading halt for 30 minutes* or until 2:30 p.m. Chicago time. If the futures contract is still limit down after 30 minutes or at 2:30 p.m., a 2-minute halt will occur, after which trading resumes.

3. Trigger: *Maximum Daily Limit*. S&P 500 futures falls 20 points from its previous close.
   Results: *No trading* can occur at a price below the maximum daily limit.

4. Trigger: *NYSE Trading Halt*. The NYSE halts trading under rule 80B.
   Results: Trading in the CME S&P 500 Index Futures contracts will be halted. Trading will be resumed only after 50 percent of the component S&P 500 stocks (measured by capitalization) have resumed trading on the NYSE.

For example, the period for which the 30-point circuit breaker on the CME is in effect depends upon whether or not the DJ 30 has fallen more than 250 points. However, a 250-point variation in the DJ 30, though unlikely, becomes more likely over time because of the natural increase in prices and in their variability.

Actual experience with circuit breakers has been limited. The major circuit breakers on the NYSE have never been triggered. Rule 80A of the NYSE and the 12-point S&P 500 circuit breaker on the CME and NYSE have been triggered, but only rarely. Thus, it is difficult to generalize about the effects of circuit breakers. Two examples are interesting.

On Friday, October 13, 1989 the stock market experienced its most significant one-day decline since the 1987 Crash. The DJ 30 dropped 190 points and the 12-point S&P 500 circuit breakers on the CME and NYSE were triggered at about 2:00 p.m. and lifted at 2:30 p.m. Then at 2:45 p.m. the 30-point S&P 500 circuit breaker in effect at that time was triggered, introducing a one-hour price floor on the CME. It does not appear that the halts in futures trading helped stabilize the stock market on that day, and the additional 105-point drop on the following Monday suggests that the circuit breaker was fighting against a drop in fundamental values.

On Monday, July 23, 1990 a sharp fall in prices in early morning trading triggered the 12-point circuit breaker and S&P 500 futures trading was stopped for 16 minutes. Resumption of trading was accompanied by a sharp rebound in the S&P 500 futures to a level that was maintained throughout the day and into the next day. Thus, in this case it appears that the circuit breaker did help restore stability.

Margin Requirements

A third set of proposals is an increase in margin requirements and the extension of those requirements to the futures markets. Two primary arguments can be made for margin requirements. The first is that they provide adequate protection against customer default for "counterparties" (brokerage firms, the Options Clearing Corporation, and futures clearing houses) by establishing higher equity requirements than the firms themselves would establish. The reason for the inability of firms to determine appropriate equity requirements is not clear, but much of the 1930s security markets legislation, which included establishing margin requirements, was predicated on the assumption that firms would choose to engage in inappropriately risky activities.

Warshawsky (1989) analyzed the maintenance margin requirements for stocks and stock index derivatives prevailing in early October of 1987. The actual margins required were compared with the margins necessary to protect the counterparty from several levels of price variation over one to five days. This study concluded that maintenance margin requirements imposed by the exchanges were adequate

The stock index option and futures markets performed appropriately during the Crash. The market that failed was the stock market, which performed poorly because of its inability to deal with the large volume of orders.

to protect the counterparty from having to make margin calls for 99 percent of one-day price movements. For individual stocks, the 25 percent maintenance margin was sufficient for 99 percent protection over five-day periods as well. Margins for options were adequate to protect against 95 percent of five-day price movements. Less protection was available, however, for the S&P 500 stock index futures: the maintenance margin of about 3.5 percent on October 16 covered 90 percent of five-day price movements. It should be noted that judging the adequacy of margin requirements in terms of not having to make a margin call is a tough standard, and it tends to understate the ability of the counterparty to avoid losses. Indeed, most margin calls are met without trouble.

The second reason for margin requirements is more relevant to the question of stock market volatility. Some traders (called "noise traders") trade on the basis of fads and fashions rather than fundamental information. Noise traders are responsible for prices deviating from fundamental values, and are the creators of bubbles and busts. Restricting their access to markets will stabilize financial markets. It seems unlikely, however, that margin requirements will screen out those with faulty information and poor methods of analysis. No relationship may exist between a trader's ability to meet margin calls and the quality of the trader's information. Indeed, it is quite possible that noise traders could be given more influence in the market if margin requirements screen out the informed traders.

Therefore, no strong evidence suggests that margin requirements are too low to protect broker-dealers and options or futures clearing houses, or that margin requirements will stabilize markets. This conclusion is buttressed by the lack of convincing evidence that margin requirements on stocks have affected the stability of the market since they were first imposed by the Federal Reserve System in 1934.

A paper by Hardevoulis (1990) found that periods of higher initial margin requirements were associated with lower stock market volatility. However, this often-cited paper has been discounted in recent work for several reasons. First, the result was almost entirely due to the inclusion of the 1930s in the sample, and does not carry over to the postwar period. Second, statistical problems with the method have been corrected and Hardevoulis' result has been rejected. On this point, see the papers by Kupiec (1989) and Salinger (1989).

V. Summary and Conclusions

The purpose of this article has been to investigate the possible reasons for, and public policy responses to, the Crash of 1987, the most prominent stock market decline experienced in several decades. This study is particularly concerned with the role played by fundamentals and market mechanisms in this event, and with the effects of recent financial innovations on the depth of the Crash.

A review of the extensive literature surrounding the Crash of 1987 reveals some important insights. The markets that have received considerable blame (the stock index options and futures markets) actually performed appropriately during the Crash: they accurately reflected market conditions and did not provide inappropriate signals to traders that would have magnified the Crash. The market that failed was the stock market, which performed poorly because of its inability to deal with the large volume of orders—a problem that exacerbated sell orders.

A corollary must be that the attention devoted to extending regulations to the stock index futures markets is misplaced. Those markets did what was expected of them, and the primary problem was elsewhere.

This study has not uncovered the "smoking gun" that would make the Crash a clearly understood phenomenon. In part, the inability to find "the" reasons for the Crash results from the fact that it was such a unique experience that it does not allow easy generalizations. In a sense, the Crash was a "hundred-year storm," a meteorological event with drastic consequences because it is of a magnitude for which protective systems are not designed, one that occurs so rarely that its ultimate causes are often poorly understood.

Observers tend to focus on the events of October of 1987, and to forget that the Crash really left very few lasting effects. Clearly, while some traders and investors were damaged, the financial system recovered rapidly with no apparent macroeconomic effects. All in all, the Crash is a tribute to the resilience of our financial markets.

References

Amihud, Yakov and Haim Mendelson. 1987. "Trading Mechanisms and Stock Returns: An Empirical Investigation." *Journal of Finance*, vol. 42, no. 3 (July), pp. 533–55.

———. 1991. "Volatility, Efficiency, and Trading: Evidence from the Japanese Stock Market." *Journal of Finance*, vol. 46, no. 5 (December), pp. 1765–89.

Bernstein, Peter. 1992. *Capital Ideas: The Improbable Origins of Modern Wall Street*. New York: The Free Press.

Black, Fischer and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, vol. 81, no. 3 (May/June), pp. 637–81.

Brady, Nicholas. 1989. *Report of the Presidential Task Force on Market Mechanisms*. January. Washington, D.C.: Government Printing Office.

Derher, William A. 1988. "Does Portfolio Insurance Ever Make Sense?" *Journal of Portfolio Management*, Summer, pp. 25–32.

Edwards, Franklin R. and Cindy W. Ma. 1992. *Futures and Options*. New York: McGraw-Hill.

Fortune, Peter. 1989. "An Assessment of Financial Market Volatility: Bills, Bonds, and Stocks." *New England Economic Review* (November/December), pp. 13–28.

———. 1991. "Stock Market Efficiency: An Autopsy?" *New England Economic Review* (March/April), pp. 17–40.

French, Kenneth and Richard Roll. 1986. "Stock Return Variances: The Arrival of Information and the Reaction of Traders." *Journal of Financial Economics*, vol. 17, pp. 5–26.

Grossman, Sanford J. 1988. "Program Trading and Market Volatility: A Report on Interday Relationships." *Financial Analysts Journal*, July–August, pp. 18–28.

Hardevoulis, Gikas. 1990. "Margin Requirements, Volatility and the Transitory Component of Stock Prices." *American Economic Review*, vol. 80, no. 4 (September), pp. 736–61.

Harris, Lawrence. 1989. "The October 1987 S&P 500 Stock-Futures Basis." *Journal of Finance*, vol. 44, no. 1 (March), pp. 77–99.

Kleidon, Allan W. and Robert E. Whaley. 1992. "One Market? Stocks, Futures, and Options During October 1987." *Journal of Finance*, vol. 47, no. 3 (July), pp. 851–77.

Kupiec, Paul H. 1989. "Initial Margin Requirements and Stock Returns Volatility: Another Look." *Journal of Financial Services Research*, vol. 3, pp. 287–301.

Mitchell, Mark and Jeremy Netter. 1989. "Triggering the 1987 Stock Market Crash: Antitakeover Provisions in the Proposed House Ways and Means Tax Bill." *Journal of Financial Economics*, vol. 24, no. 1 (September), pp. 37–68.

O'Brien, Thomas J. 1988. "The Mechanics of Portfolio Insurance." *Journal of Portfolio Management*, vol. 14, no. 3 (Spring), pp. 40–7.

Roll, Richard. 1988. "The International Crash of October, 1987." *Financial Analysts Journal*, September/October, pp. 19–35.

Rubinstein, Mark. 1985. "Alternative Paths to Portfolio Insurance." *Financial Analysts Journal*, July/August, pp. 42–52.

Rubenstein, Mark and Hayne Leland. 1981. "Replicating Options with Positions in Stock and Cash." *Financial Analysts Journal*, July/August, pp. 42–52.

Salinger, Michael A. 1989. "Stock Market Margin Requirements and Volatility: Implications for Regulation of Stock Index Futures." *Journal of Financial Services Research*, vol. 3, pp. 121–38.

Santoni, Gary J. 1988. "The October Crash: Some Evidence on the Cascade Theory." Federal Bank of St. Louis *Review* (May/June), pp. 18–33.

Securities and Exchange Commission. 1988. *The October 1987 Market Break*. Washington, D.C.: Government Printing Office. February.

Seyhun, H. Nejat. 1990. "Overreaction or Fundamentals: Some Lessons from Insiders' Response to the Market Crash of 1987." *Journal of Finance*, vol. 45, no. 3 (December), 1363–88.

Sofianos, George. 1988. "Margin Requirements on Equity Instruments." Federal Reserve Bank of New York *Quarterly Review* (Summer), pp. 47–60.

Teweles, Richard J. and Edward S. Bradley. 1987. *The Stock Market*. Fifth Edition. New York: John Wiley and Sons.

Warshawsky, Mark. 1989. "The Adequacy and Consistency of Margin Requirements in the Markets for Stocks and Derivative Products." Staff Study 158, Board of Governors of the Federal Reserve System, Washington D.C., September.

Equity in School Finance: State Aid to Local Schools in New England

Katharine L. Bradbury

Assistant Vice President and Economist, Federal Reserve Bank of Boston.
Andrew Evans, Catherine Jew, David Mann, and Thomas Miles provided able research assistance. The author is grateful to Lynn Browne, Steve Gold, Jeannette Hargroves, Yolanda Kodrzycki, and Bob Tannenwald for helpful comments, and to the Departments of Education in the six New England states for providing data.

Disparities in school spending have been the key school finance problem for decades. Americans have a long tradition of universally available public education, thought to play an important role in "leveling the playing field" by providing equal educational opportunity to rich and poor alike. However, because education is provided locally and because localities differ in both resources and preferences, expenditures on schooling vary noticeably from one area to another. Most states provide considerable funding to local school districts and attempt to use these funds to equalize spending, but disparities in spending between rich and poor districts remain sizable within many states. In Rhode Island, for example, the richest district spends about twice as much per pupil as the poorest district.

Spending disparities have not disappeared, despite considerable attention and a fair degree of consensus on the ideal of equal educational opportunity, for several reasons. First, state-provided or state-mandated uniform schooling runs counter to another tenet of U.S. public education, local decisionmaking. Second, various factors muddy the equating of equal educational opportunity with equal dollars.

Against a background of recent school finance trends, this paper discusses concepts of equity in school finance—the trade-off between equality and local control and the difficulties involved in using spending as an indicator of educational opportunity (Part II). Parts III and IV lay out the types of school aid formulas that states employ to promote such equity and the factors that influence the way local districts' final spending decisions respond to aid. Part V summarizes the key issues in recent school finance court cases. The remaining sections of the paper then describe current school aid programs in the New England states, examining intrastate patterns of school spending and local tax rates to evaluate the aid's equalizing effects.

I. Recent Trends in School Finance

Public elementary and secondary schools in the United States have historically been operated at the local level, by cities, towns, counties, and independent local school districts. State governments, however, have long had an important role in *financing* local schools, as well as setting standards or otherwise regulating local school systems. Table 1 shows the growing role of state governments in financing elementary and secondary education in the United States. State government funding surpassed the local share in the 1970s, and states now finance almost half of K–12 school spending nationwide.

The state role varies widely across the states, however. Hawaii's public schools are operated by state government; at the other end of the spectrum, New Hampshire's local public schools receive less than 10 percent of their funds from the state (Table 2). While New Hampshire is extreme, all the New England states except Maine have retained an above-average *local* role in revenue-raising for schools.

States have provided more funds to local school districts over the years in large part to foster inter-district equity by offsetting the unequal distribution of local resources. (For school districts in most states, property taxes are the major—if not the only—local resource.) In providing funds, most states have attempted to ensure an adequate education for all public-school children, but they have gone about it in a variety of ways. Furthermore, a number of state legislatures have revised their aid formulas in response to court decisions, sometimes repeatedly. As discussed in part V below, such decisions have typically found that existing aid programs did not provide a sufficient offset to the spending disparities that result from local districts' dependence on the property tax, and thereby ran counter to state constitutional provisions requiring universal access to "adequate" or "equitable" public schooling. Court decisions requiring school finance reform picked up in the 1990s "in a flurry of activity not seen since the early 1970s" (Kosterlitz 1990), and a number of additional cases, including one in Massachusetts, are currently being heard.

The difficulty state governments face is that a remedy is not obvious. When local districts control the final spending decision, the policy tool available to state government is school aid; states can design their aid formulas to provide incentives for districts to behave in desired ways, but no formula can guarantee a specific outcome. The next two sections outline the conceptual issues involved in school finance, both goals and trade-offs, and explain the major types of school aid formulas states have adopted to attain these goals.

Table 1
Public School District Revenues by Source of Funds
Percent of Total

School Year	Federal	State	Local
1939–40	1.8	30.3	68.0
1949–50	2.9	39.8	57.3
1959–60	4.4	39.1	56.5
1969–70	8.0	39.9	52.1
1979–80	9.8	46.8	43.4
1989–90	6.1	47.2	46.6

Source: U.S. Department of Education, National Center for Education Statistics, *Digest of Education Statistics 1992*, Table 148.

II. Standards of Equity in School Finance

The trade-off between local control and equal opportunity frames the school finance debate. One view is that any disparities in education are bad—no child should receive a poorer-quality education than any other child, even if that outcome reflects the relative priorities that taxpayers in each district place on schools versus other public and private purchases they could make. But in adhering to another American value, local decisionmaking regarding the education of local children, states implicitly accept and even endorse educational disparities attributable to different preferences. Thus, most states appear to have a somewhat looser definition of equity, not that schools should be equal, but that educational disparities should not primarily reflect disparities in residents' wealth (as indicated by the property tax base) or income levels. When an association between wealth and per pupil spending persists even when state school aid is equalizing, states directly confront the need for other, controversial strategies that generally involve some reduction in local sovereignty, such as requiring all districts to spend above some minimum or even capping allowable levels of spending.

The second major difficulty in "solving" the school finance problem is that most discussion and measurement focus on spending, but the basic issue

Table 2
Sources of Revenue for Public Elementary and Secondary Schools, 1989–90
Percent of Total

	Federal	State	Local
United States	6.1	47.2	46.6
Alabama	11.2	60.0	28.8
Alaska	12.8	62.4	24.8
Arizona	7.9	43.5	48.6
Arkansas	9.6	56.8	33.6
California	6.6	66.9	26.5
Colorado	4.8	38.1	57.1
Connecticut	4.6	43.1	52.3
Delaware	7.3	66.8	25.9
District of Columbia	9.8	—	90.2
Florida	6.2	51.2	42.5
Georgia	6.3	53.1	40.5
Hawaii	10.1	87.3	2.6
Idaho	8.0	60.2	31.8
Illinois	5.9	32.8	61.3
Indiana	4.9	57.7	37.4
Iowa	4.9	49.1	46.0
Kansas	5.0	44.2	50.9
Kentucky	9.8	68.5	21.6
Louisiana	10.1	55.5	34.4
Maine	5.4	53.1	41.4
Maryland	4.6	37.7	57.7
Massachusetts	4.7	34.5	60.8
Michigan	5.7	26.8	67.4
Minnesota	4.1	52.4	43.5
Mississippi	15.5	56.2	28.3
Missouri	5.5	40.0	54.4
Montana	9.0	45.9	45.1
Nebraska	5.9	23.1	71.0
Nevada	4.2	38.0	57.8
New Hampshire	2.8	8.4	88.8
New Jersey	3.8	39.8	56.4
New Mexico	12.3	72.9	14.8
New York	5.1	40.7	54.1
North Carolina	6.4	66.8	26.8
North Dakota	9.8	44.8	45.5
Ohio	5.4	43.6	51.1
Oklahoma	5.6	57.0	37.4
Oregon	6.1	25.1	68.8
Pennsylvania	5.2	43.6	51.2
Rhode Island	4.9	43.1	52.0
South Carolina	8.0	50.0	41.9
South Dakota	11.5	25.9	62.6
Tennessee	9.0	45.8	45.2
Texas	7.3	41.9	50.8
Utah	6.6	56.6	36.8
Vermont	4.3	32.2	63.4
Virginia	5.3	33.1	61.7
Washington	5.8	71.6	22.6
West Virginia	7.5	65.7	26.8
Wisconsin	4.1	40.2	55.7
Wyoming	5.0	51.2	43.8

Source: U.S. Department of Education, National Center for Education Statistics, *Digest of Education Statistics 1992*, Table 149.

is equality of educational inputs or outcomes across rich and poor districts, not simply equality of expenditures. Giving all students "the same quality" education may require spending more money in one district than another, if the children have different educational needs or if the cost of educational inputs varies significantly from one district to the next. Thus many states attempt to make more equal not spending per se but spending adjusted to reflect differences in needs or costs.[1] Or they may equalize specific physical inputs (such as teachers per student).

Per pupil spending or inputs (or even outcomes) that are equal *across* districts do not ensure that each school or each student *within* a district fares equally well. Nevertheless, on practical grounds—the status of current state school aid programs and the availability of data—this article focuses on dollars per student at the district level.

Furthermore, even equal need- or cost-adjusted dollars will not result in the same education everywhere because the way school funds are spent varies considerably across districts and states. For example, some districts spend more on curriculum development, some on experienced teachers' salaries, some on science labs. In recent years, the equity issue has become entangled in broader educational reform issues relating to the overall quality of elementary and secondary education in the United States, as doubts are raised about American workers' ability to compete in international markets. The competitiveness debates bring to the fore disagreements among parents, teachers, school administrators, other education professionals, and business observers regarding what to teach and how best to teach it to elementary and secondary school students, even when resources are ample. Real public school spending per pupil has risen at a faster pace than most government spending over the past few decades, yet many observers believe that the overall quality of education has not risen concomitantly.

Money alone cannot equalize educational opportunity or attain the nation's education goals. Recent education reform efforts focus on performance assessment and accountability for students, teachers, schools, and school districts. To address these issues, states attach strings to their school aid, from setting

[1] This poses considerable measurement problems because local cost indexes are not readily available and student needs (for special, vocational, or bilingual education, for example) are difficult to measure and even more difficult to translate into spending requirements.

standards to explicitly tying aid amounts to performance. Ultimately, states pursuing the goal of equal educational opportunity might attempt to ensure that the educational *outcomes* for children in different school districts were not associated with the income or wealth of the district. But measurement of student outcomes is impossible without consensus on what students should know and tests that measure what students do know at the beginning and end of school.[2] Alternatively, advocates of school choice argue that the "market test" implicit in school choice schemes (within a public school district, among districts, or among public and private schools) holds schools and school systems fully accountable, rewarding excellence and eliminating poor performers. While systems of interdistrict school choice can, in theory, give all students equal access to all schools statewide, making moot the issue of unequal resources among districts, in practice, access to the "best" schools is necessarily limited at any point in time.

To date, few states have incorporated such school performance standards or accountability into the design and implementation of their equalizing school aid formulas; most continue to puzzle over how to integrate educational (or "structural" or "systemic") reform initiatives with school finance reform. Kentucky is an example of a state that revamped the governance, structure, and program of its schools, as well as financing, after the state Supreme Court found "the whole gamut of the [state's public] school system" unconstitutional; the new system focuses on student performance. The most difficult challenge in accountability schemes, including school choice, is to develop approaches that reward the progress made by schools or districts (and not simply attainment levels) so as to avoid leaving some districts' students behind in a self-reinforcing downward spiral, as poor performance leads to reduced aid and fewer resources lead to performance declines.

In addition to equality of educational opportunity for public school students, states often have a second criterion for equity in school finance, that taxpayers in different districts should not face unequal tax burdens to provide equal education to their children. While the local-control philosophy honors the educational *choices* districts may make, this fiscal neutrality standard seeks to remove the wealth-related constraints that would bias such choices in poor versus rich districts. Thus, taxpayer equity is viewed as an intermediate step toward equalizing spending—given local decisionmaking, districts facing similar tax rates are thought likely to choose similar spending levels. However, many states have found themselves back in court after implementing a new school aid plan because they did not get beyond that intermediate step: legislators have found to their chagrin that most equalizing aid programs are more successful in bringing school tax rates closer together than in reducing spending disparities.

III. School Aid Formulas

The two key decisions states must make regarding aid to school districts are the total aid budget and the formula that determines how much aid each district will receive. States' equalizing aid formulas generally fall into one of three categories: foundation plans, guaranteed tax base or percentage-equalizing programs, and combination plans.[3] These formulas are equalizing in the sense that school districts with fewer local resources receive more aid than richer districts, but they differ regarding whether the funding for each district is invariant to actual district spending or *matches* local school spending.

A foundation program provides aid to each district in proportion to the number of students and in inverse proportion to the local property tax base per pupil (or other measure of local resources). A district's actual school spending does not affect its aid amount. Over three-quarters of the states use a foundation-type program for their basic school support. (See Gold et al. 1992, Table 4, p. 18.)

[2] See, for example, Downes (1992) who finds that school finance reform in California in response to Proposition 13 and the *Serrano* court decision reduced spending disparities across districts but not disparities in student performance (sixth-grade test scores). He argues that the performance disparities are partly attributable to faster-rising costs of educating the students in poorer districts as well as to actions taken by richer districts to raise more school revenue outside of state-imposed revenue limits.

[3] A few states provide basic aid through flat grants that simply provide the same dollar amount per pupil to every school district. But these grants are not considered equalizing because they fail to provide more aid per pupil to poorer districts. Many states also have a variety of K–12 education objectives *that are not equity-related* which they attempt to address through school aid programs of a "categorical" nature. Categorical aid reflects the specific programmatic priorities that state government wants to support at the local level. States vary considerably in the emphasis they place on categorical versus "basic equalizing" aid. (See Gold et al. 1990, Table 7, p. 38.) Some states use mandates instead of categorical aid to impose their priorities on local school districts. Indeed, all states impose some educational standards on local school districts. Categorical aids and mandates are beyond the purview of this article.

"Pure" foundation formulas revolve around a statewide "foundation" spending level per pupil and a target tax rate. The per-pupil aid to any one district is calculated to make up the difference between the foundation spending level and the amount of revenue per pupil that the district could raise by applying the target tax rate to its local tax base. (See Appendix A for a mathematical formulation of the key features of each type of grant formula.) As a condition for receipt of aid, pure foundation plans require that each district actually spend at least its foundation amount (not just the aid amount) on schools, but some states do not impose this minimum spending requirement.

Guaranteed tax base (or percentage-equalizing) plans match the dollars that districts spend, with the matching rate varying inversely with local resources. In a pure guaranteed tax base program, the per-pupil aid to any district makes up the difference between what the district actually raises and what a district that had the guaranteed tax base per pupil would raise using the given district's actual tax rate. That is, regardless of the actual size of the local tax base, all districts have access to the same school tax base per pupil; levying a given tax rate yields the same per pupil revenue (taxes and aid combined) in a rich or poor district.[4]

Both types of grant can be adjusted for cost differences among districts if a state decides to provide more money per pupil to higher-cost or "needier" students. The most common adjustment is made by substituting "weighted" students for simple student counts in the formula, where the weights reflect the proportional increase in cost thought to be associated with educating students in designated categories relative to a "regular day elementary" student. Some states also adjust for interdistrict cost differences, if the costs of standard inputs (wage rates or materials prices) vary across districts and can be measured at the district level.

States must also decide whether to interfere with the operation of the formula in order to limit year-to-year changes in aid for individual districts. Some states limit the "maximum loss" that any district can incur in a year; others guarantee districts the previous year's aid amount and run the formula only to allocate each year's aid increment. (This latter strategy obviously is feasible only when the total aid pool is growing.) Over time, such restrictions can significantly reduce the equalizing impact of any aid program.

Either type of formula can be made more or less equalizing by changing the size of the state's total aid budget or altering the formula's parameters. An increase in total state dollars, distributed through any equalizing formula, is likely to lead to more equal spending across districts. So are actions such as raising the foundation amount and lowering the target tax rate, or raising the guaranteed tax base. These actions alter the number of districts that qualify for aid and the degree to which aid funds are focused on poorer districts.

Having set the program parameters, each state must also decide how to treat districts that, according to the formula, should receive negative aid amounts because they have "ample" resources. Theoretically, the state could "recapture" the negative aid amounts but, more commonly, those richer districts simply receive zero equalizing aid or a minimum dollar amount per pupil. Indeed, the equalization goal is often compromised slightly in order to "buy" all districts into the plan with minimum aid.[5] The compromise is that the richest districts are able to spend more at any tax rate or tax themselves more lightly to support any spending level than poorer districts, which would not be possible if the state recaptured funds from them. Figure 1 shows graphically the tax rates required to attain certain spending levels under foundation or guaranteed tax base plans with *no* recapture. For the richest districts seen toward the right side of the figure,[6] tax rates begin to decline with increments to the tax base.

While guaranteed tax base programs reward lo-

[4] The percentage-equalizing label is also applied to other formulas that share the key characteristics of the guaranteed tax base program: the state reimburses a fraction of actual school expenditures, and the fraction varies inversely with the local district's ability to raise revenue. That is, poor districts receive aid equal to a larger fraction of their actual spending than rich districts.

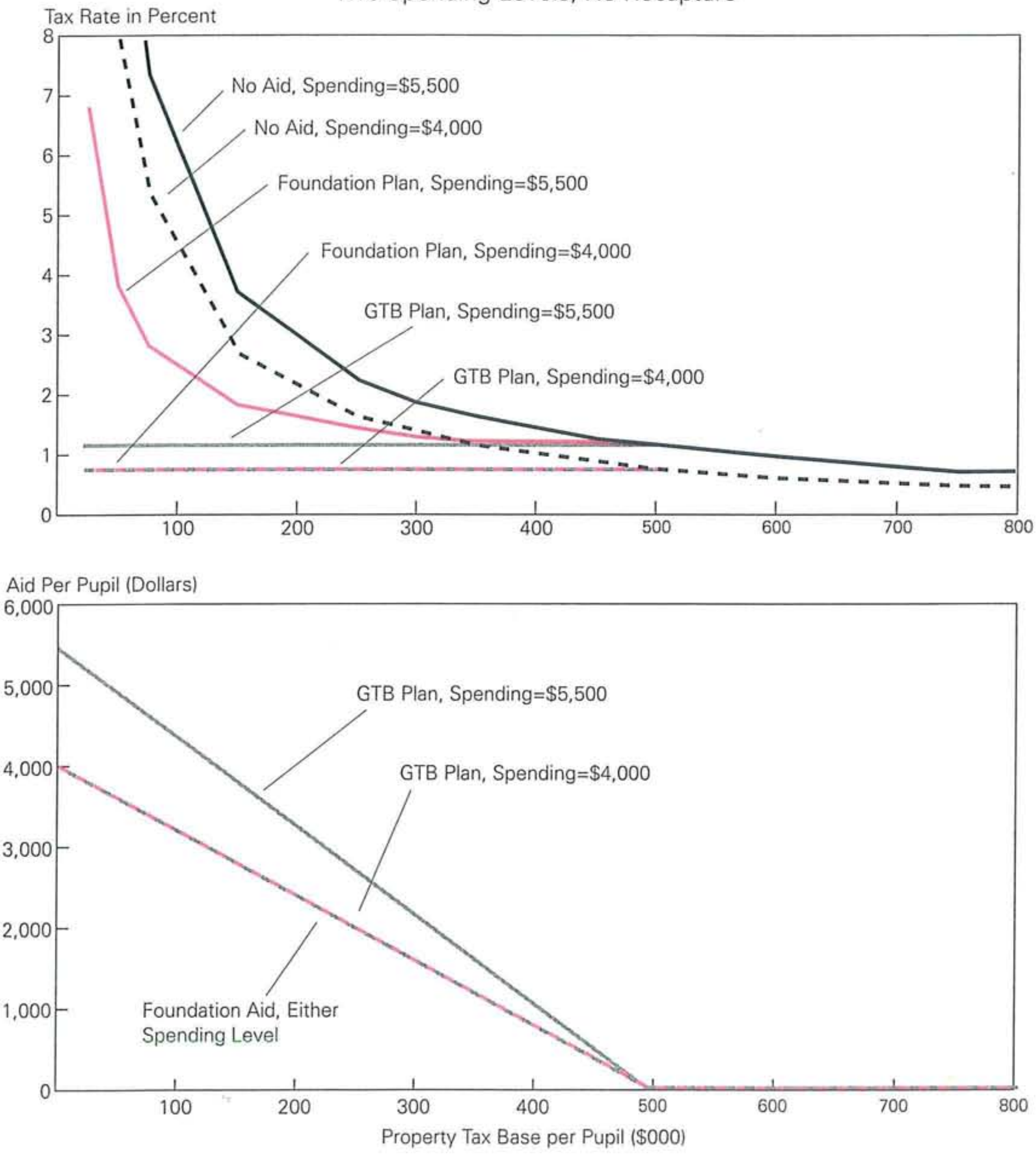
[5] Political feasibility aside, recapture may violate some state constitutions. A Wisconsin court overturned its recapture provision and Maine voters repealed a recapture provision after a lawsuit charged that it was unconstitutional. In Montana, recapture was viewed as functioning as a state tax and was upheld. (See LaMorte 1989, p. 12.)

[6] Districts with tax bases in excess of the guaranteed tax base have access to their richer base. Similarly, in a foundation plan, districts with tax bases in excess of the foundation breakeven (the base at which the target tax rate yields the foundation amount per pupil) are able to collect property taxes in excess of foundation spending levels with a tax rate below the target rate. The sample formulas depicted in Figure 1 were chosen to make the two plans as similar as possible: the foundation is $4,000 per pupil, the target tax rate is 0.8 percent, and the guaranteed tax base is $500,000 per pupil—the tax base that would yield foundation spending at the target tax rate.

Figure 1

Foundation Plan, Guaranteed Tax Base Plan, or No Aid

Two Spending Levels, No Recapture



Note: Foundation= $4,000, target tax rate=0.8%, guaranteed tax base (GTB)=$500,000.

cal "effort" so that any local tax rate yields the same aid-plus-tax revenue in a rich or poor district, under a foundation program poor districts adding spending above the foundation amount face a steeper increase in their tax rates than do rich districts adding the same per pupil spending. (Compare the two programs' tax rates required for the higher spending level in Figure 1.) To remedy this drawback of foundation programs but still retain a required minimum spending level, states have enacted combination plans, adding a guaranteed tax base program to "level the playing field" beyond the foundation level of spending. If the foundation spending level is low, the guaranteed tax base add-on is especially important, since many districts will want to spend more than the foundation amount and poorer districts would face a significant tax rate disadvantage with the foundation plan alone.

IV. Equalizing School Aid and Local Spending Decisions

The essence of local decisionmaking is that local school districts make the final decision about how much money to spend per pupil, presumably taking into consideration the amount of aid provided by the state as well as their own ability to raise revenue. The dollars provided by state aid raise the total resources available to a school district, and specific provisions of the aid formula may further alter the spending incentives faced by the district. But, depending on local preferences and other factors,[7] those incremental resources might be used by the community for additional school spending, other local public spending, or to reduce taxes.

Earmarking—the requirement that school aid add, dollar for dollar, to school spending—is controversial because it treads the line between state authority and local decisionmaking. Furthermore, it is almost impossible to enforce. Most states would like local districts to decide how much local money to spend on schools in the absence of aid and then add total school aid to the local money. But it is virtually impossible for state officials providing aid to know what the district would otherwise have spent from local sources. In most cases, the only enforceable requirement is that school spending (from all sources) must equal or exceed the amount of school aid; if a district would spend more than that regardless of aid, the requirement does not alter school spending.[8]

As noted earlier, poor districts typically spend less per pupil than rich districts and pay higher school tax rates in the absence of aid. If poor districts get more aid dollars per pupil than rich districts, then presumably both the spending disparities and the tax rate disparities will fall when aid is provided or augmented. The degree to which the funds add to spending or reduce taxes, of course, depends on the parameters of the aid program chosen by the state as well as patterns of spending, tax rates, tax bases, and voter preferences in the absence of aid.

Because they match locally raised funds, percentage-equalizing or guaranteed tax base programs reduce the effective tax rate needed to obtain each additional dollar of school funds; economists call this a change in the "price" a district pays for school spending. Indeed, the basic rationale for guaranteed tax base programs is to equalize the tax price of school funds across districts, ensuring that poorer districts do not face higher tax rates than rich districts to raise the same dollars per pupil. Economists have analyzed the effect of additional resources ("income effect") and tax-price changes ("price effect") on local governments' spending decisions. The consensus from studies of the income effect seems to be that an additional non-matching dollar of school aid raises school spending by about 50 cents (Odden and Picus 1992, pp. 85–86); adding the price effect, the spending response would presumably be greater to aid dollars distributed through matching programs (percentage-equalizing or guaranteed tax base plans). (See Figure 2 for a simple graphic representation of how the two types of aid affect the choices open to local districts.) Thus for any given pool of aid funds, guaranteed tax base or percentage-equalizing plans are likely to be more successful at equalizing spending—if the matching rate is considerably higher for poor than for rich districts—than are foundation plans.[9] Conversely, since one element of foundation plans is the target tax rate, they are generally more successful at reducing tax rate disparities, at least among districts spending at or below the foundation amount.

A guaranteed tax base plan removes the relative tax rate *dis*incentive to spend additional dollars on schools faced by poorer districts. But even so, poor districts may still choose to spend less on schools than rich districts, for several reasons. First, ability to pay may not be fully captured in a tax-base measure: given the same per-pupil tax base, taxpayers in a high-income district may be willing to spend more on schools than low-income taxpayers for whom any

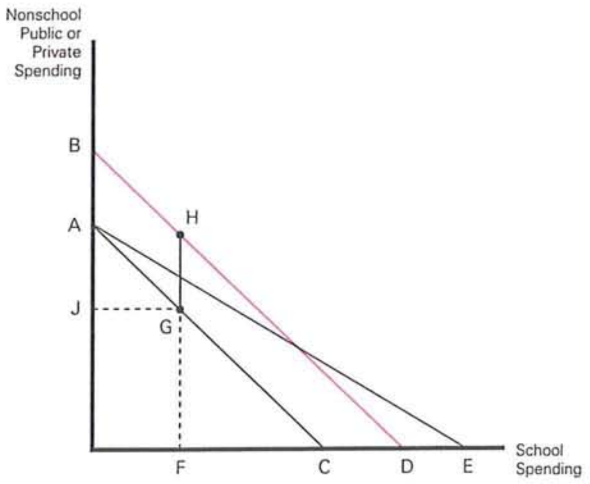
[7] Local responses also depend on the institutional structure through which local spending decisions are made (type of government, presence of binding tax limitations, referendum requirements, and the like) and on differences in the perceived price of school spending that are independent of the aid formula, such as voter perceptions of the community's ability to export taxes (which might depend on the fraction of property that is nonresidential).

[8] If a district's school budget would otherwise be shrinking by more than the incremental aid amount, then strict enforcement of maintenance of effort provisions (requiring locally raised funds to equal or exceed the previous year's amount) could force each additional dollar of aid to add to spending in the short run.

[9] This will not be the case if richer districts are much more sensitive to matching aid than are poor districts.

Figure 2

Schematic Diagram of Price Shift vs. Income Shift



AC represents a district's budget and spending options in the absence of aid (A=C). Along AC, the district's voters decide how to divide its total resources between school and nonschool (including private consumption) spending. For example, a district at point G on AC is spending F on schools and J on nonschool items.

A guaranteed tax base or percentage-equalizing aid program shifts possibilities to AE, providing no aid if school spending is zero (at point A) and increasing amounts of aid the higher is school spending. By contrast, a foundation plan shifts the possibilities to points along BD (or AGHD if spending greater than the foundation F is a requirement for receipt of aid). BD is parallel to AC because the amount of foundation aid is invariant to actual spending.

For poorer districts, local resources (measured by point A, which is equal to C) are smaller. Both types of formula offer more aid to poorer districts, given any spending choice. That is, for a poorer district, the distance between C and D and between C and E would be greater than that shown above for a district with average resources.

given property tax bill represents a greater burden relative to income. Second, some poor districts, especially big-city districts, may choose to spend less per pupil than rich districts on account of greater competing demands for other (nonschool) government spending (sometimes referred to as "municipal overburden").

Instead of trying to use additional aid to induce poor districts to add to spending, states sometimes simply overrule local choices by adopting a foundation plan with required minimum spending. For maximum equalizing impact, a state would adopt a combination plan with a fairly high required foundation, and add on a high guaranteed tax base for above-foundation spending.

The choice of a school aid formula and funding level by state policymakers thus reflects the nature of a state's concern for interdistrict disparities and its commitment to reducing them. After a discussion of school finance court cases in states around the country, the subsequent two sections describe the size and form of school aid programs in each of the New England states and current patterns of spending and tax rate disparities among local school districts in each state.

V. School Finance Court Cases

School finance suits have been brought in over half the states, based on evidence that the existing system of school finance, with local schools supported by a combination of local tax revenues (raised from unequal property tax bases) and aid from the state government, provides considerably fewer resources to educate students living in the state's poorest districts. Depending on the wording of the education clause of each state's constitution, these suits allege that such a situation is not "equitable" or "efficient" or fails to provide equal protection or an "adequate" education for residents of poorer districts.

Court decisions in about half the cases to date have overturned the state's system of school finance and about half have upheld it (Odden and Picus 1992, p. 36). The courts have generally avoided prescribing a remedy, preferring to leave that to elected state legislators. Indeed, some courts that have failed to find their state's school finance system unconstitutional have nevertheless explicitly criticized the status quo, but deferred to legislators to initiate reforms. In several states, the courts have been asked to take a second or third look after state legislatures have enacted equalizing reforms, and have found the reformed system still unconstitutional. (For example, see the box for a summary of ongoing Texas attempts at reform, which have recently taken an unusual turn.)

One of the key differences between courts that have overturned and upheld state school finance

A New Approach in Texas

The Texas Legislature voted this February to amend the state constitution to "force wealthy school districts to transfer money to poor ones" (Verhover 1993). The amendment would allow the Legislature to shift 2.75 percent of all state and local school revenue from districts with high property wealth to poor ones. Legislators approved the amendment as superior to the two alternatives, consolidation of the state's 1,000-plus districts into regional units or a court-ordered cutoff of financing that would shut down the schools several weeks before the scheduled end of the school year. If the amendment passes, Texas will be the first state to put a "recapture" provision into its constitution.

This amendment (which goes before the voters on May 1) is the most recent action in a 25-year struggle to insure fairer spending on the public schools in Texas. In an early case, filed in 1968 and decided in 1973 (*Rodriguez v. San Antonio Independent School District* 337 F. Supp. 280 (W.D. Tex. 1971); rev'd, 411 U.S. 1 (1973)), the U.S. Supreme Court upheld the constitutionality of Texas's system of school finance. But judging the appeal of a 1984 suit (*Edgewood Independent School District, et al. v. Kirby, et al.* 77 S.W. 2d 391 (1989)), the Texas Supreme Court in 1989 found the system to be in violation of the state constitution. A substantially reformed school finance system was found to be still unconstitutional by the state district court in 1990, a decision upheld by the state Supreme Court in 1991.

The Texas constitution requires the legislature to "make suitable provision for the support and maintenance of an efficient system of public free schools." The state courts have interpreted "efficiency" as a loose fiscal neutrality standard, requiring that each school district would have "substantially equal access to similar revenues per pupil at similar levels of tax effort" (Walker 1990, p. 11). The post-reform system found to be still unconstitutional in 1990 included a foundation plan plus a percentage-equalizing add-on for spending above the foundation. One ongoing complaint has been insufficient state funding of this system. Another sticking point in each court decision was the "inefficiency" of concentrated resources in "property-rich districts that tax low" (Walker 1990, p. 10), which helps explain why the constitutional amendment for redistribution and a regional consolidation plan are considered to be alternative remedies.

systems is the emphasis they place on local control. Local control is often the key rationale for upholding the status quo. By contrast, courts overturning school finance systems have concluded that the existing system provided meaningful local control only to wealthy districts or that the state constitution put a higher priority on equity than on local control (LaMorte 1989).

Although evidence is typically offered regarding wide disparities among districts in *spending*, most of the state court decisions have been based on a fiscal neutrality standard, that the current system of financing does not provide all districts *equal access* to school revenues. For example, in the case of *Serrano v. Priest*, the California Supreme Court decided that inequities in educational opportunities existed under California's foundation plan since "two school districts levying the same tax rate but with different taxable wealth per pupil would have different per-pupil spending" (quotation from Downes 1992). As discussed earlier, a guaranteed tax base or percentage-equalizing state aid program can offset the unequal distribution of property tax bases to give poor and rich districts *access* to the same revenues for the same tax effort, while allowing poor districts to channel some of their aid into tax relief as well as higher spending. Some decisions have combined fiscal neutrality with concern for the "adequacy" of support for basic education; in these cases, a foundation (or combination) plan would be more useful in bringing all districts above an explicit minimum spending level.

Despite the focus on equal *ability* to raise revenue (fiscal neutrality), some courts have also indicated that spending disparities beyond a certain range are unacceptable. For example, the 1971 *Serrano* decision stated that wealth-related spending differences among school districts in California could not exceed $100 per pupil (a limit that was later adjusted for inflation). A New Jersey court decree required the

Table 3
State Aid to Local Public Schools in the New England States, 1990–91

	Connecticut	Maine	Massachusetts	New Hampshire[a]	Rhode Island	Vermont
State Aid per Pupil ($)	2,633	2,436	1,775	280	2,442	1,521
State Aid Relative to Total School Budget (percent)	40.0	49.6	31.5	6.1	40.5	31.2

Note: These figures are pupil-weighted averages of school district data for each state.
[a]New Hampshire data are for 1989–90.
Source: Calculations based on data supplied by Departments of Education in the New England states. See Appendix B for definition of aid included for each state.

state to raise the spending level of the poorest 28 districts to that of the wealthiest suburban districts. But according to Odden and Picus (1992), "only Wyoming has created a standard of equal expenditure per pupil that the school finance system must meet" (p. 36). Since additional school aid dollars can go only so far in raising spending in poorer districts, such limits on spending disparities have forced states such as California and New Jersey to consider capping spending in richer districts.

VI. State Aid to Education in New England

The New England states, like states elsewhere in the nation, approach local school aid in a variety of ways. Only Rhode Island has a percentage-equalizing aid program; the other five states have foundation-type plans (some with a percentage-equalizing second tier). The states also vary widely in the size of their commitment to local public schools: the state government in New Hampshire provides less than $300 of public school aid per pupil, on average, while all the other New England states provide more than $1,500 per pupil (Table 3).

Table 4 summarizes the key characteristics of the basic equalizing aid programs in the New England states (and Appendix A describes each state's formula). The paragraphs that follow highlight the key characteristics for each state.

Connecticut uses a foundation-type formula for the bulk of its aid, providing more dollars per pupil than any other New England state. The foundation spending level changes annually in line with statewide school spending. The (implicit) target tax rate is the foundation divided by "state-guaranteed wealth," defined as some multiple of median town wealth. The legislature can and does alter the equalizing impact of the program by changing the multiple. "Wealth" is a composite measure of ability to pay, reflecting both per capita income and the size of the local property tax base.

Spending at or above the foundation per "need student" is virtually required; districts face a steep aid penalty if they spend less. The "need student" count gives heavier weight to low-income students and students scoring below the remedial level in a mastery test. (Appendix A describes pupil "weights.")

Maine's basic school aid is officially a foundation plan for operating costs with a small percentage-equalizing "quality incentive adjustment" add-on that partly matches higher spending for poorer districts. However, because districts spending less than the foundation amount have their aid reduced proportionally, the formula actually operates like a guaranteed tax base plan in which the guaranteed tax base is equal to the foundation divided by the target tax rate. Thus the formula has a matching aspect both below and above the foundation spending level, with the matching rate lower for above-foundation spending.[10]

The foundation spending levels (different for K–8 and high school) are set on the basis of average statewide spending in the previous year, adjusted for cost increases. The (implicit) target tax rate is approximately the statewide average tax rate that would be required to raise 45 percent of the statewide foundation, because the state covers an average of 55 percent of school costs[11]—a considerably higher level of state participation than in the other New England states.

[10] The state also provides a smaller amount of aid through a percentage-equalizing program for school debt service costs and "special program" costs (including special needs and vocational education).

[11] This 55 percent figure does not exactly match Maine's data in Table 3 because of different concepts and measures. The formula refers to 55 percent of state-local spending, disregarding federal funds. And Table 3 reports pupil-weighted averages of school district data, not statewide averages.

Table 4
Key Characteristics of Basic Equalizing Aid Formulas in the New England States

	CT	ME	MA	NH	RI	VT
Type of Formula						
Foundation	X		X	X[a]		X
Guaranteed tax base		X[b]			X	X[c]
Minimum Spending Required	X[d]					
Student Weights[e] Reflect						
Grade level		X[f]	X[g]	X	X[g]	X
Poverty or income	X		X			X
Special ed./vocational ed./bilingual			X	X		
Separate Aid Program Available for						
Special ed./vocational ed./bilingual	X	X		X	X	X
Some Districts Receive						
Zero basic aid				X		X
Flat minimum aid[h]	X	X	X		X	X
Limited year-to-year changes[i]	X[j]	X	X[k]		[m]	X
Ability to Pay Includes						
Property tax base	X	X	X[n]	X	X	X
Resident income	X			X	X	X
Basic Equalizing Aid as Percent of Total Aid	62	64	55	60	88	67

[a]School tax rate and locally raised revenues also enter the formula.
[b]Basic aid is officially foundation-type with an add-on for spending above the foundation level. But because aid is reduced proportionally if a district spends below the foundation amount, Maine's plan operates like a GTB formula with the matching rate halved for spending above foundation.
[c]Percentage-equalizing add-on for half of spending above foundation level.
[d]If a town fails to spend required amount, aid is reduced by twice the shortfall.
[e]See Appendix A for complete description of weights.
[f]No official weights, but different elementary and secondary foundation amounts are equivalent to weighting by grade level.
[g]Among grade levels, only kindergarten students are weighted (because they attend only part of the day).
[h]Flat minimum aid (or zero aid) is usually provided to the richest districts.
[i]Changes may be limited by "hold harmless" or "maximum loss" provisions, or by other restrictions on reductions in aid from one year to the next.
[j]After the 1990–93 transition, minimum and enhancement aid require aid growth.
[k]Formula is applied only to incremental aid each year.
[m]Only temporarily in transition years.
[n]Property tax base is expressed per capita, not per pupil.
Source: Program descriptions supplied by Departments of Education in the New England states and Gold et al. (1992).

Massachusetts has an official equalizing school aid formula, but during the 1980s another general-purpose aid formula determined the (school plus municipal) aid total,[12] and the school aid formula was used simply to label a fraction of those additional assistance funds as school aid. The aid data shown in Table 3 (and elsewhere) reflect only the official school formula amounts. Both formulas operated on an incremental basis; that is, each community received the aid it had received the prior year, and the formula determined only the increment to aid each year. Furthermore, in fiscal years 1990 through 1992, neither aid formula was operative; municipal and school aid *cuts* were apportioned partly in proportion to total aid dollars and partly on a per capita basis. School aid was increased for fiscal year 1993, mostly on a per-pupil basis.

The official formula for schools is "Chapter 70," a foundation-type program using weighted pupils and the per capita property tax base as the measure of local ability to pay.[13] The implicit target tax rate is

[12] The general-purpose aid for municipal governments reflects school-related costs (a multiple of weighted pupils) as well as other municipal costs, and provides aid in proportion to the measured gap between local needs (costs) and local revenue-raising capacity. This "additional assistance" formula can be decomposed into school and nonschool components; the school component operates much like a foundation plan.

[13] The use of a per capita rather than a per pupil measure of the property tax base was designed to recognize "municipal overburden" concerns and increase slightly the amount of aid going to the (mostly urban) communities with relatively few school-age children per capita.

equal to a statewide constant[14] multiplied by the tax rate that would be required for a community with the statewide average tax base per capita to raise the foundation amount per pupil.

New Hampshire has a "foundation" plan in the broad sense: the state provides money in proportion to weighted pupils and in inverse proportion to ability to pay. The formula's measure of ability to pay reflects not only property tax base per weighted pupil, but also per capita income and local tax effort relative to state averages. The most notable characteristic of New Hampshire's aid program is its small size: for a district with average wealth, income, and tax effort, aid would be equal to 8 percent of the per-pupil foundation amount (set equal to statewide average spending) times the number of weighted pupils, where the weights reflect special needs students and their placement and high school, vocational, and regular elementary student counts. If funds are insufficient (as they have been in the last few years), the percentage of each district's foundation amount covered by aid is reduced by a constant percentage point amount. Districts are not required to spend the foundation amount to qualify for aid.

Rhode Island's basic aid matches actual expenditures, with the matching rate varying inversely with ability to pay. According to the formula, the state should reimburse about 50 percent of the average district's spending,[15] although when funding falls short (as it did in FY93), aid to each district is reduced proportionally. Legislation guarantees that no district's aid will cover less than 15 percent of spending in fiscal year 1993 and 9 percent thereafter (down from a 28.5 percent minimum in earlier years). The guaranteed tax base implicit in the formula varies across communities; for a community with family income at the statewide median, it is roughly equal to twice the statewide per-pupil property tax base.

Vermont has a classic foundation plan with a percentage-equalizing add-on for above-foundation spending, and tiered treatment of richer communities for which the formula would yield negative aid.[16] Pupil weights reflect high school and elementary student counts, poverty, and transported students by density (Appendix A). The foundation tax rate varies to make the aid more equalizing than a standard foundation plan; it is adjusted for each district in proportion to local income.[17]

Year-to-year declines in aid are limited to $100 per pupil by a "maximum loss" provision.[18] Districts are not required to spend at or above the foundation level, but the state also provides "supplemental aid for above-average expenditures," which uses each district's ratio of basic aid to foundation cost as a matching rate for one-half of spending above the foundation level.

VII. Patterns of School Spending in New England

Table 5 reports measures of disparity in school spending and tax bases across school districts within the six New England states. It indicates that disparities in operating spending among school districts are greatest in Massachusetts and least noticeable in Rhode Island.[19] As noted earlier, however, spending disparities can reflect different local priorities; it is those associated with resource differences that arouse concern.

Depending on which dispersion measure is used, disparities in property wealth may be greatest in Maine (according to the restricted range ratio), Vermont (coefficient of variation), Rhode Island (McLoone index)[20], or Connecticut (gini coefficient).

[14] The constant is an adjustment factor used to make the formula distribution consistent with the aid budget.

[15] As is the case for Maine, this figure does not match Rhode Island's data in Table 3 for definitional and measurement reasons. See footnote 11.

[16] If a district has resources in excess of need, then basic aid is zero, but if resources are 1 to 1.5 times need, then the district qualifies for "minimum aid" of $150 per pupil. Districts with resources that exceed 1.5 times need are "gold" towns and receive no aid.

[17] Income is measured as adjusted gross income per exemption indexed to the statewide average. For each unit of the index, the tax rate (in mills) is adjusted either up or down by 27 cents (so if income is twice the state average, the district's foundation tax rate will be 27 cents above the base rate). The statewide foundation tax rate was 1.085 percent in FY91 and 1.1175 percent in FY92.

[18] The number of communities "on" the formula (not "gold" zero aid, not minimum aid, not maximum loss) declined from 186 in 1988 to 91 in 1992, while the number of "maximum loss" towns grew from 34 to 99.

[19] Rhode Island has many fewer school districts than the other New England states, which may tend to reduce the measured interdistrict disparities.

[20] The coefficient of variation and the gini coefficient for the property tax base per pupil are lowest in Rhode Island, indicating low disparity, while the McLoone index suggests Rhode Island has the highest wealth disparity. The McLoone index reflects data on only half the districts—those with values below the median—indicating how far below the median they are on a pupil-weighted basis. The disagreement among the three measures presumably indicates that wealth inequalities in Rhode Island are concentrated below the median; that is, the districts above the median are fairly closely clustered, but some districts have considerably lower per-pupil wealth.

Table 5
Measures of Interdistrict Disparity within Each New England State, 1990–91

Measure	Connecticut	Maine	Massachusetts	New Hampshire[a]	Rhode Island	Vermont
Number of Districts	169	258	347	153	35	251
Number of Students	464,056	212,262	834,216	166,752	133,666	97,465
Operating Spending per Pupil ($)						
Average[b]	6,816	4,165	5,082	4,378	6,112	4,240
Minimum	5,051	2,264	3,251	2,583	5,104	2,321
Maximum	9,823	9,669	8,861	8,067	10,738	6,935
Restricted Range Ratio[c]	1.50	1.49	1.77	1.74	1.41	1.67
Coefficient of Variation	12.6	13.3	19.8	15.7	8.0	16.7
McLoone Index[d]	.930	.940	.901	.971	.953	.919
Gini Coefficient[e]	.070	.070	.107	.093	.040	.087
Property Tax Base per Pupil[f] ($000)						
Average[b]	552.7	262.9	512.3	439.6	176.9	361.4
Minimum	160.4	27.0	171.0	101.7	50.8	117.9
Maximum	2,620.5	5,856.5	7,039.0	5,235.8	1,513.1	17,171.3
Restricted Range Ratio[c]	5.39	6.67	4.85	3.33	3.82	5.25
Coefficient of Variation	67.2	72.0	60.9	88.1	46.5	106.0
McLoone Index[d]	.776	.687	.700	.808	.678	.747
Gini Coefficient[e]	.306	.217	.229	.211	.157	.219

[a]New Hampshire data are for 1989–90.
[b]Pupil-weighted.
[c]Ratio of value at 95th percentile to value at 5th percentile, pupil-weighted.
[d]An index of equity measuring ratio of sum of variable for pupils in districts below pupil-weighted median to what sum would be if they were at median; lower values indicate more inequality.
[e]The gini coefficient measures how unequally the object (spending or tax base) is distributed among pupils—higher values indicate more inequality.
[f]Property tax base is equalized; the data are states' estimates of the market value of taxable property in each district.
Source: See Table 3. See Appendix B for definition of operating spending used for each state.

Two of the measures indicate that wealth disparities are least pronounced in New Hampshire and the other two point to Rhode Island. In all the states and according to all the reported measures, disparities in per-pupil property wealth are considerably greater than spending disparities. Hence some equalizing is occurring in all six states.

In simplified terms, per-pupil spending in any district depends on its tax base per pupil, the school tax rate applied to that tax base, and the amount of school aid it receives.[21] Tax base disparities would translate directly into spending disparities if all districts levied the same school tax rate and aid were not equalizing. But the typical pattern is that poor districts tax themselves more heavily in order to raise their spending somewhat closer to that of rich districts. This means that school aid must operate on two fronts to achieve fiscal neutrality; it must provide enough funds to bring spending per pupil in the poorest districts up to that in the richest districts and, furthermore, bring school tax rates in the poorest districts down to those in the richest districts, on average.[22]

Table 6 narrows the focus from disparities per se to fiscal neutrality, quantifying the degree to which tax bases, spending, tax rates, and aid in the richest districts differ from their counterparts in the poorest districts. The table ranks the school districts in each state by property tax base per pupil, groups them into five equal groups (quintiles) from poorest to richest, and reports the average values of key variables for each group.

[21] This ignores local revenues other than the property tax and school aid from the federal government, but these are very small revenue sources compared with property taxes and state school aid.

[22] That is, more equalizing aid (or smaller underlying tax base disparities) will likely result in more equal tax rates as well as more equal spending. However, for a given aid distribution and pattern of wealth disparities, a trade-off between the two dimensions of fiscal neutrality exists: the more heavily poor districts tax themselves, the more equal spending will be.

Table 6
District Characteristics by Property Tax Wealth per Pupil, 1990–91
Average Values in Wealth Quintiles for School Districts in the New England States

State and Characteristic	Overall Average	Poorest Quintile	Second Quintile	Middle Quintile	Fourth Quintile	Richest Quintile	Ratio: Richest/Poorest
Connecticut							
Property Tax Base per Pupil[a]	560.1	251.9	345.7	448.7	640.8	1,110.0	4.4
Operating Spending per Pupil ($)	6,710	6,179	6,361	6,375	6,958	7,669	1.2
School Tax Rate (mills)	9.19	10.85	10.01	9.67	8.60	6.83	.6
State Aid per Pupil ($)	2,478	3,906	3,195	2,391	1,897	997	.3
State Aid/School Budget (%)	38.5	63.5	50.3	37.9	27.5	27.5	.4
Maine							
Property Tax Base per Pupil[a]	402.6	94.5	144.8	219.3	352.5	1,211.6	12.8
Operating Spending per Pupil ($)	4,152	3,853	4,039	3,864	4,114	4,900	1.3
School Tax Rate (mills)	9.02	10.93	10.03	9.29	9.02	5.79	.5
State Aid per Pupil ($)	2,373	3,561	3,380	2,603	1,741	570	.2
State Aid/School Budget (%)	46.2	72.0	64.9	52.8	32.4	8.7	.1
Massachusetts							
Property Tax Base per Pupil[a]	624.8	249.2	344.6	442.8	596.3	1,484.2	6.0
Operating Spending per Pupil ($)	5,009	4,258	4,528	4,765	5,365	6,124	1.4
School Tax Rate (mills)	7.39	7.77	8.21	8.12	7.71	5.16	.7
State Aid per Pupil ($)	1,680	2,628	2,164	1,590	1,070	949	.4
State Aid/School Budget (%)	29.8	50.9	40.1	28.0	17.4	12.4	.2
New Hampshire[b]							
Property Tax Base per Pupil[a]	573.4	241.3	327.9	392.7	513.5	1,407.7	5.8
Operating Spending per Pupil ($)	4,387	3,817	3,815	4,535	4,392	5,389	1.4
School Tax Rate (mills)	12.01	14.49	14.26	13.74	11.10	6.39	.4
State Aid per Pupil ($)	314	937	358	177	104	4	...
State Aid/School Budget (%)	6.7	21.7	7.2	3.2	1.4	.1	...
Rhode Island							
Property Tax Base per Pupil[a]	245.7	98.8	157.4	190.5	224.6	557.2	5.6
Operating Spending per Pupil ($)	6,262	5,699	6,060	6,034	6,194	7,321	1.3
School Tax Rate (mills)	6.88	7.75	6.94	7.13	7.39	5.20	.7
State Aid per Pupil ($)	2,289	2,992	2,671	2,250	1,909	1,622	.5
State Aid/School Budget (%)	37.5	52.8	44.0	37.3	30.8	22.5	.4
Vermont							
Property Tax Base per Pupil[a]	535.2	177.5	236.1	306.0	429.4	1,531.6	8.6
Operating Spending per Pupil ($)	4,187	3,842	4,001	4,179	4,216	4,700	1.2
School Tax Rate (mills)	10.74	10.96	12.01	12.45	11.91	6.33	.6
State Aid per Pupil ($)	1,403	2,877	2,120	1,408	576	33	...
State Aid/School Budget (%)	28.6	60.2	43.7	27.8	10.8	.5	...

Note: Figures in overall average column above do not match averages in Tables 3 and 5 because these data are district-weighted, and data in Tables 3 and 5 are pupil-weighted.
...Less than 0.05.
[a]Property tax base is in thousands of dollars of equalized value; quintiles are defined in terms of this variable.
[b]New Hampshire data are for 1989–90.
Source: See Table 3. See Appendix B for definitions of operating spending and state aid for each state.

On a quintile-to-quintile basis, Table 6 indicates that tax base disparities are highest in Maine and second highest in Vermont (the richest quintile has 13 times the average tax base per pupil of the poorest quintile in Maine, and nine times in Vermont). Spending per pupil, however is only about 30 percent higher in the richest quintile than in the poorest quintile in Maine, and 20 percent higher in Vermont. Spending disparities among wealth quintiles are smaller than wealth disparities partly because school aid is quite equalizing (the richest quintile averages only one-fifth the aid per pupil of the poorest quintile in Maine and less than one-fiftieth in Vermont) and partly because the poorer communities tax themselves more heavily (the school tax rate is half as high in the richest quintile as in the poorest in Maine and about 60 percent as high in Vermont).[23] That is, spending disparities in Maine and Vermont are substantial, but not as wide as the underlying tax base disparities because of the offsetting effects of more aid and higher tax rates in poorer districts.

Spending differences between rich and poor districts in Rhode Island are similar to those in Maine and Vermont[24] (the richest quintile spends 1.3 times as much as the poorest), but the underlying property tax base disparities are considerably smaller than in Maine and Vermont. Rhode Island's smaller wealth disparities do not translate into noticeably smaller spending disparities because state aid is less equalizing (the richest quintile averages one-half the aid per pupil of the poorest quintile), and differences in tax effort are also less pronounced (the richest districts average about 70 percent of the school tax rates of the poorest).

Connecticut also enjoys less pronounced property tax base disparities among wealth quintiles than the other states, and its school aid is not as concentrated on poorer communities. The smaller wealth disparities translate into lower spending disparities (as in Vermont, spending per pupil is only 20 percent higher in the richest quintile than in the poorest). Connecticut's school aid is less equalizing than Maine's and Vermont's, and its tax rate disparities are smaller (the richest quintile's tax effort is about 60 percent as high as the poorest's).

Massachusetts shows mid-range tax base disparities between quintiles, but the highest spending disparities (spending per pupil is over 40 percent higher in the richest quintile of districts than in the poorest). The data suggest that spending disparities between rich and poor districts persist in Massachusetts more than in Maine and Vermont because the distribution of aid is less equalizing in Massachusetts—the richest quintile of districts receives about 40 percent as much per-pupil aid as the poorest quintile—and because tax rates are more similar across quintiles, which means that tax base disparities translate more directly into school revenue disparities. Local property taxes in Massachusetts (for school and municipal purposes combined) are constrained by Proposition 2½, a property tax limitation measure enacted in 1980.

In New Hampshire, like Massachusetts, spending was 40 percent higher in the richest quintile than in the poorest, on average, although its tax base disparities are relatively small. As in Vermont, school aid is highly focused on poor districts, but school aid in New Hampshire comprises such a small portion of the school budget that it cannot offset a very large fraction of the wealth disparity, even in the poorest districts. What keeps spending disparities from being even larger in New Hampshire is tax rate disparities; school tax rates in the poorest districts are more than twice as high as in the richest districts, on average.

In sum, spending disparities between rich and poor districts are largest in Massachusetts and New Hampshire and smallest in Connecticut and Vermont. Of the latter two, Vermont deserves special attention since the underlying tax base differences are greater there than in Connecticut, but school aid helped bring spending disparities down. Table 6 also indicates that tax rates in rich and poor districts are most disparate in New Hampshire and most similar in Massachusetts and Rhode Island.

VIII. Current Challenges and Issues

One interpretation of the data in Table 6 might focus on the remarkable similarity among the New England states of spending disparities across districts, given the wide variation in underlying tax base disparities, school aid, and school tax rates. This similarity suggests a "comfort" level, common among the New England states, beyond which state policymakers (who determine aid) and local decisionmakers (who determine school taxes) find spending disparities unacceptable. For example, the Rhode Island pattern in Table 6, compared with the other New England states, might occur because Rhode Island state legislators perceive less need for its aid program to be equalizing since its wealth disparities are less pronounced, and because Rhode Island's local school district officials need not levy widely varying tax rates to raise similar amounts of revenue.

[23] Note that three states, most notably Vermont, show higher average school tax rates in some of the mid-wealth quintiles than in the poorest.

[24] The disparity measures in Table 5 indicated that Rhode Island's spending disparities were lower than those in the other New England states, while Massachusetts' were highest. By contrast, Table 6 suggests spending differences between rich and poor districts are smallest in Vermont and Connecticut, with Rhode Island and Maine falling in the middle tier. This discrepancy between the two tables indicates that some of the spending disparities in Vermont and Connecticut are associated with preference or cost differences or other nonwealth factors.

However, that "comfort level," if it exists, is not at the point of equal spending in rich and poor districts. In all the New England states, students and parents living in poorer districts typically experience lower spending and higher school tax rates than their counterparts in richer districts. Even in the region's most fiscally neutral states, the richest one-fifth of the districts spend 20 percent more per pupil than the poorest one-fifth, on average, and these spending disparities are larger when the very richest districts—the top 10 percent, for example—are compared with the very poorest.

The patterns of spending, wealth, aid, and tax rates examined in this article suggest that some strategies are likely to be more successful than others in achieving greater fiscal neutrality—greater equality of spending and tax rates across school districts—in the New England states. As just noted, Vermont has relatively low spending disparities despite fairly pronounced tax base disparities—aid goes a long way toward equalization. This provides some endorsement for the "combination" type of aid formula Vermont uses, with a percentage-equalizing add-on to a foundation plan. The foundation plan has no minimum spending requirement and no recapture, but does provide zero aid to the richest towns.[25] With this type of formula, Vermont could achieve greater equality in spending, should it wish to do so, by adding more dollars to the school aid pool; its aid commitment is currently in the low to middle range among the New England states.

Maine, like Vermont, has a small percentage-equalizing add-on to its aid formula for the poorest districts, and makes a more sizable contribution to school aid than most of the New England states. While spending at or above the foundation is not required, aid is reduced proportionally if spending falls below the foundation level. This aid program is quite equalizing; it brings spending disparities to the middle range, despite the largest tax base disparities in the region. Tax rate disparities, however, remain sizable in Maine, probably because the aid formula is matching, which focuses its influence on spending.

Connecticut's spending disparities are also low, largely because the underlying tax base disparities are relatively small. Consequently, Connecticut can achieve as much equalization as the other states with less targeting of aid to the poorest districts. Perhaps equally important, its requirement that districts spend at least the foundation amount brings all districts' spending above that minimum (while foundation aid minimizes the tax rate impacts of such a requirement).

New Hampshire's spending disparities between rich and poor districts are relatively high, and poor communities tax themselves relatively heavily. The state's aid program is quite equalizing as far as it goes, but it does not provide much money, even to the poorest districts. A substantially greater commitment of state funds to school aid would be required to achieve fiscal neutrality comparable to other states,[26] but New Hampshire's current budget difficulties have prompted a move in the other direction—reductions in aid and even some "trial balloon" proposals that the state remove itself from school standard-setting.

Rhode Island, alone among the New England states, provides school aid that explicitly matches local school spending, with a higher matching rate for poorer districts. Starting with small tax base disparities, Rhode Island displays a low level of tax rate disparities but mid-level spending disparities. Rhode Island's aid dollars are not more equalizing, largely because a fairly high minimum matching rate (aid covered no less than 28.5 percent of spending in fiscal year 1991) gives rich districts almost as much incentive as poor ones to increase spending. The state has passed legislation to reduce the statutory minimum matching rate to 9 percent after fiscal year 1993, which should enhance the aid's equalizing impact. A combination plan with a binding foundation minimum (and a percentage-equalizing program as an add-on for above-foundation spending) might be more successful than the current formula in raising school spending in the state's poorest districts.

The impact of Massachusetts' school aid is difficult to assess, since the dollars labeled as school aid (and hence measured in this and other studies) bear little relationship to the actual aid dollars provided to each community for combined school and municipal purposes each year. In terms of outcomes, the aid

---

[25] The growing number of communities receiving "maximum loss" aid, however, blunts the equalizing impact of the aid program. See footnote 18.

[26] In theory, the state government's lack of any broad-based taxes and the local governments' access to the property tax might provide an argument for adoption of an expanded foundation plan (or a guaranteed tax base plan) *with recapture*; that is, aid to poorer districts would be augmented by redistributing property tax funds from richer districts. In practice, however, the property tax is already heavily used in New Hampshire (because it is the only broad-based tax), and recapture plans have been politically palatable in few other states.

The Massachusetts School Finance Case

Arguments in the case of *McDuffy v. Robertson* began in February 1993 before the Supreme Judicial Court of Massachusetts. Begun as *Webby v. King* in 1978, the case was moved to the back burner in the mid 1980s after the enactment of new aid targeted to districts spending less than 85 percent of the statewide average per pupil. New plaintiffs replaced the original "Webby" students in 1990 when cuts in school aid exacerbated the alleged failure of school finance reforms of the 1980s to produce equitable spending patterns.

The constitution of the Commonwealth requires that the state "cherish" education. As a result of this vague language, the government is arguing that it has no legal obligation to fund or otherwise ensure the adequacy or equity of education provided by local governments. To its credit, according to this view, the state has already taken steps to equalize funds through school aid, reducing but not eliminating interdistrict disparities.

By contrast, the plaintiffs, children from some of the state's poorest districts, charge that "the system by which Massachusetts finances the construction and operation of elementary and secondary public schools . . . denies the plaintiff students an adequate education and equal educational opportunities and advantages solely because they reside in cities and towns with relatively low real property wealth."

As in many states, the court case has altered the school finance reform climate in Massachusetts even before the Court's ruling. Both sides agree that remedies should emanate from the legislature, but they have very different opinions regarding how well reform measures currently being considered measure up.

does not appear to be as equalizing as other states', since spending disparities are relatively high. The incremental nature of aid in Massachusetts may limit the degree of equalization: the formula (for school or general purpose aid) applies only to the new dollars added to the aid pool each year; furthermore, fiscal year 1991 represented the second year of aid cuts to Massachusetts cities and towns. By contrast with spending disparities, tax rate disparities are relatively low, perhaps because of Proposition 2½, the state's property tax limitation measure. Indeed, Prop 2½ may contribute to school spending disparities by limiting the degree to which poor communities can tax themselves more heavily in order to raise revenues comparable to their wealthier neighbors. (See the box for a discussion of the most recent Massachusetts court case on school finance.)

Massachusetts added money to its school aid pool last year, and plans to do the same this spring for fiscal year 1994 aid. The FY93 increment was distributed mostly on a per-pupil basis, but consideration was given to a proposal for a combination formula, with spending required at or above a reasonably adequate foundation level and percentage-equalizing aid for spending above the foundation.[27] This year's proposals include some of the same features. More money explicitly for schools combined with a phase-out of the incremental approach and a shift toward a combination formula would undoubtedly enhance the program's equalizing effects.

All told, the New England states could further reduce spending disparities among school districts by moving toward combination plans, with a percentage-equalizing add-on to a foundation formula such as Vermont uses. Making the foundation a binding minimum at an adequate spending level brings up spending in the lowest-spending districts. Putting more money into the aid pool makes possible greater equalization or allows pursuit of general education support without sacrificing equalization; such budget decisions, of course, must also consider other priority uses for state funds. Straightening out these school finance issues is a critical challenge as the states face key ongoing decisions regarding broader educational reform and accountability.

[27] The Massachusetts Business Alliance for Education proposal included a foundation spending level of $5,000 per pupil with a percentage-equalizing add-on to allow poorer districts to spend above the foundation without facing markedly higher school tax rates.

Appendix A: Equalizing School Aid Formulas

I. Generic Formulas

Foundation Plan

$APP_i = F - (t \times PBPP_i)$; aid makes up the difference between the foundation spending level and what the local tax base yields at the target tax rate.

(And $APP_i = 0$ if $(t \times PBPP_i) > F$ when "negative aid" is not recaptured.)

(Furthermore, $APP_i = 0$ if $EPP_i < F_i$ when aid is contingent on districts spending at least the foundation amount per pupil.)

Then, assuming spending per pupil equals aid plus local school revenue from the property tax (a simplifying assumption that ignores federal aid, other state aid, and non-property tax local revenue),

$r_i = t + [(EPP_i - F)/PBPP_i]$ = target tax rate plus spending above (or below if allowed) the foundation divided by per-pupil tax base;

$DSA_i = APP_i \times n_i = (F \times n_i) - (t \times PBPP_i \times n_i)$ = the foundation amount times the number of pupils minus what the target tax rate raises with the local tax base;

where

APP_i is aid per pupil to school district i,
r_i is district i's school tax rate,
EPP_i is district i's spending per pupil,
$PBPP_i$ is district i's property tax base per pupil,
DSA_i is district i's total school aid,
n_i is the number of pupils in district i,
F is the statewide foundation amount per pupil, and
t is the statewide target tax rate.
Note that subscript i refers to district i and unsubscripted variables are statewide constants.

The "foundation" label is also applied to other formulas in which aid is proportional to the number of pupils (n_i) and negatively related to the per pupil property tax base ($PBPP_i$) or other measure of district resources.

Guaranteed Tax Base or Percentage-Equalizing Plan

$APP_i = r_i \times (GTB - PBPP_i)$; aid makes up the difference between what the local tax base yields and what the guaranteed tax base would yield at district i's tax rate.

(And $APP_i = 0$ if $PBPP_i > GTB$ when "negative aid" is not recaptured.)

Then, assuming spending per pupil equals aid plus local school revenue from the property tax,

$r_i = EPP_i/GTB$ = spending divided by the guaranteed tax base;

$DSA_i = APP_i \times n_i = r_i \times [(GTB \times n_i) - (PBPP_i \times n_i)]$ = the local tax rate times the difference between the guaranteed tax base and the actual tax base;

where the above definitions apply, and

GTB is the guaranteed tax base, a statewide constant.

The "percentage equalizing" label is also applied to other formulas in which aid is matching (aid is proportional to actual spending) and the matching rate varies inversely with the local per pupil property tax base.

For example,

$APP_i = (1 - LS_i) \times EPP_i$, where LS_i is the fraction of spending that district i must raise locally,

$LS_i = c \times PBPP_i/PBPP$, the local share is a constant c times the size of district i's property tax base per pupil relative to the statewide average property wealth per pupil. Then

$APP_i = (1 - c \times PBPP_i/PBPP) \times EPP_i$; aid covers a certain fraction of total spending, where the fraction varies inversely with the relative size of local property wealth per pupil.

Then, assuming spending per pupil equals aid plus local school revenue from the property tax,

$r_i = c \times EPP_i/PBPP$

$DSA_i = APP_i \times n_i = (1 - c \times PBPP_i/PBPP) \times EPP_i \times n_i$, a fraction of total spending

$= r_i \times \{(PBPP/c \times n_i) - (PBPP_i \times n_i)\}$, the local tax rate multiplied by the difference between [some multiple of the statewide average tax base per pupil times district i's pupils] and district i's tax base.

Thus, under the usual simplifying assumptions, this percentage-equalizing formula is equivalent to the GTB formula, where the implicit GTB = PBPP/c.

Combination Plan

$$APP_i = \begin{cases} F - (t \times PBPP_i) & \text{if } EPP_i \le F; \\ & \text{like foundation plan for spending at or below foundation level.} \\ & \text{(And } APP_i = 0 \text{ if } (t \times PBPP_i) > F, \text{ when "negative aid" is not recaptured.)} \\ F - (t \times RBPP_i) + (r_i - t) \times (GTB - PBPP_i) = & \\ F - (t \times GTB) + r_i \times (GTB - PBPP_i) & \text{if } EPP_i > F; \\ & \text{foundation aid plus guaranteed tax base plan for spending above foundation level.} \end{cases}$$

Then, assuming spending per pupil equals aid plus local school revenue from the property tax,

$$r_i = \begin{cases} t + [(EPP_i - F)/PBPP_i] & \text{if } EPP_i \le F, \text{ the target tax rate, or less if below-foundation spending is allowed.} \\ t + (EPP_i - F)/GTB & \text{if } EPP_i > F, \text{ the target tax rate plus spending in excess of the foundation divided by the guaranteed tax base.} \end{cases}$$

$$DSA_i = \begin{cases} (F \times n_i) - (t \times PBPP_i \times n_i) & \text{if } EPP_i \le F \\ (F \times n_i) - (t \times GTB \times n_i) + r_i \\ \quad \times [(GTB \times n_i) - (PBPP_i \times n_i)] & \text{if } EPP_i > F \end{cases}$$

Incorporating differential costs or needs into the formulas

Weighted pupils (w_i) can be substituted for the simple pupil count (n_i) if the state decides to provide more money per pupil to higher-cost or "needier" students. The weights reflect the proportional increase in cost thought to be associated with each type of student, and hence the proportional increase in aid provided by the state.

II. Equalizing Aid Formulas in the New England States

The equalizing aid formulas used by the New England states can be expressed in a form comparable to the "generic" aid formulas above. The abbreviations used above have the same meanings below. Subscript i refers to school district i; unsubscripted variables are statewide constants.

Connecticut

$$DSA_i = (F \times w_i) - (F/SGW) \times (PCY_i/PCY^{highest}) \times PBPW_i \times w_i$$

where w is weighted students, SGW is the state guaranteed wealth, PCY is per capita income, and PBPW is property tax base per weighted pupil. In this context, the "target tax rate" is F/SGW, that is, the rate required for a community with state-guaranteed wealth to raise the foundation amount per pupil. Local ability to pay is measured as property tax base per weighted pupil adjusted for per capita income relative to the highest town's per capita income.

Maine

For a community spending the foundation amount or more,

$$DBA_i = [1 - \{(OCMR \times PBPP_i)/F\}] \times [(F^e \times n_i^e) + (F^s \times n_i^s)] = (F \times w_i) - (OCMR \times PBPP_i \times w_i) \quad \text{if } EPW_i \geq F$$

where DBA_i is basic equalizing aid to district i, OCMR is the statewide "operating cost mill rate" (a target tax rate), the superscripts e and s refer to elementary and secondary students, the weights are derived from relative elementary and secondary per pupil "foundation" amounts (that is, the weights are F^e/F and F^s/F), and EPW_i is school spending per weighted pupil.

When a district does not raise its share of foundation spending, then aid is reduced proportionally:

$$DSA_i = (r_i/OCMR) \times \{(F \times n_i) - (OCMR \times PBPP_i \times n_i)\} \quad \text{if } EPW_i < F$$

where r_i is the local property tax rate. Thus the adjustment is in proportion to the ratio of the district's actual tax rate to the target tax rate. Rearranging terms,

$$DSA_i = r_i \times \{(F/OCMR \times n_i) - (PBPP_i \times n_i)\} \quad \text{if } EPW_i < F,$$

which is the same as a guaranteed tax base formula in which the guaranteed tax base is equal to the foundation breakeven (implicit GTB=F/OCMR). Maine also provides "quality incentive adjustment" aid for districts receiving nonzero basic aid that spent above the foundation two years prior. The state matches one-half of that "excess" spending, and the matching rate is the ratio of basic aid to foundation spending:

$$DAA_i = .5 \times \{1 - (OCMR \times PBPP_i/F)\} \times (EPW_i \times w_i - F \times w_i) \quad \text{if } EPW_i > F.$$

For districts that qualify for both basic aid and the quality incentive adjustment, total aid is

$$DSA_i = DBA_i + DAA_i.$$

Massachusetts

$$DSA_i = \{F \times w_i\} - \{P \times (F/PBPC) \times PBPC_i \times w_i\}$$

$$= \{F \times w_i\} - \{P \times (F/PBPW) \times [(w_i/pop_i)/(W/POP)] \times (PBPW_i \times w_i)\}$$

where P is a statewide constant, PBPC is property tax base per capita, PBPW is the property tax base per weighted pupil, POP is population, and unsubscripted capital letters refer to statewide averages. In this context, the target tax rate is $P \times \{(F/PBPC)\}$, a multiple of the tax rate that would be required for a community with the statewide average tax base per capita to raise the foundation amount per pupil. Ability to pay is expressed in per capita terms; that is, it is adjusted for the ratio of weighted school pupils to population.

New Hampshire

$$DSA_i = .08 \times (PBPW/PBPW_i) \times (PCY/pcy_i) \times .5\{[(STR_i/STR) \times (PCY/PCY_i)] + (SR/SR_i)\} \times w_i \times F$$

where STR is the effective school tax rate, SR is locally-raised school revenues per weighted pupil, and unsubscripted capital letters refer to statewide averages. This formula cannot be expressed in the same form as the typical foundation plan because the local property tax base is in the denominator rather than the numerator, but as in any foundation plan, aid is inversely related to the local property tax base per weighted pupil and proportional to the number of weighted pupils.

Rhode Island

$$DSA_i = (1 - \{.5 \times ([EWAV_i/n_i]/[EWAV_s/n_s])\}) \times EPP_i$$

where EWAV is the property tax base adjusted to reflect median family income (MFI below) and the subscript s refers to statewide totals. If reimbursable expenditures were equal to aid plus local revenues ($r_i \times PBPP_i \times n_i$) and if $EWAV_i = (PBPP_i \times n_i) \times (MFI_i/MFI)$, then

$$DSA_i = r_i \times \{[2 \times (EWAV_s/n_s) \times (MFI_i/MFI) \times n_i] - (PBPP_i \times n_i)\}.$$

The implicit guaranteed tax base thus varies across communities in proportion to relative income; for a community with family income at the statewide median, the guaranteed tax base is roughly equal to twice the statewide property tax base per pupil.

Vermont

$$DBA_i = F \times w_i - (t + x_i) \times (PBPW_i \times w_i)$$

where DBA is a district's basic school aid and x is an additive change in the foundation ("target") tax rate proportional to local income.

Vermont also provides supplemental aid for districts that qualify for nonzero basic aid and spend above the foundation; half of expenditures in excess of the foundation are matched, and the matching rate is the ratio of basic aid to foundation cost:

$$DAA_i = .5 \times \{1 - (t + x_i) \times (PBPW_i/F)\} \times (EPWi - F) \quad \text{if } EPW_i > F$$

where DAA_i is supplemental aid and EPW_i is spending per weighted pupil; supplemental aid is zero if

$$\{PBPW_i \times (t + x_i)\} > F.$$

For a district with spending above the foundation and property tax base low enough to qualify for basic aid, total equalizing aid is

$$DSA_i = DBA_i + DAA_i.$$

III. Pupil Weights Used in New England School Aid Formulas

Connecticut uses "need students" in its formula. The count of need students gives heavier weight to low-income students (1.25 for AFDC) and students scoring below the remedial level in a mastery test (1.25). "The mastery count is a proxy measure of the number of students testing below the remedial standard on the mastery tests plus a bonus for improvements in mastery test scores."

Maine has no pupil weights, but does have a higher foundation per-pupil operating rate for high school than for K–8 students, based on actual statewide spending patterns. In FY91, the high school foundation was 41 percent higher (= 4213/2982) than the elementary school foundation. Other weights are not used in the general purpose "operating cost" aid formula because Maine has separate "program cost" aid for early childhood education, special education, vocational education, transportation operations, and bus purchases.

Massachusetts uses pupil weights in its aid formulas (both Chapter 70 and additional assistance). Weight is 1.0 for regular day students and higher for special education (4.0), bilingual (1.4), occupational day (2.0), and residential (4.0) students; low-income (AFDC) students receive an additional weight of 0.2. Pre-kindergarten and kindergarten students are weighted 0.5 (or 0.7 if bilingual) because they attend only half-day.

New Hampshire also incorporates weights into its formula, based on state average expenditures per pupil for eight educational programs. Weights are 1.0 for regular elementary student, and higher for regular high school (1.21), high school vocational education (2.01), special education in-district special education classroom (2.57), special education mainstreamed (2.12), out-of-district special education day (7.08), out-of-district special education residential (8.72), special education preschool day placement (3.37).

Rhode Island—no weights.

Vermont's weights reflect elementary (1.0) and secondary school (1.25 for grades 7–12) student counts, poverty (1.25 for children in families receiving food stamps), and transported students by density. The transported student weights are higher for transported students in more sparsely settled catchment areas: sparsely settled (3 or fewer students transported per square mile) = 1.0714, moderately settled (over 3 but no more than 9) = 1.05, and densely settled (more than 9) = 1.0385.

Appendix B: Definitions of Key School Measures for the New England States

Connecticut

Operating spending defined as "Net Current Expenditures"—"Those expenditures on behalf of public elementary and secondary education from all sources: state, local, federal, and other." Excluded from NCE are debt service and capital outlay, reimbursable regular education transportation, adult education, and expenditures on behalf of nonpublic schools.

State school aid defined as "Equalized grants"—total state equalized formula aid, the sum of state grants that encompass an equalized distribution formula. In 1990–91, these included education cost sharing (the foundation formula), special education-regular, public and nonpublic transportation, school construction, adult education, health services, and vocational education equipment/OIC.

Pupil count is average daily membership. Observations are cities and towns; data for regional school districts have been attributed back to member cities and towns.

Maine

Operating spending defined as "Operating Costs Residential"—all general fund costs reported in each School Administrative Unit's annual Financial Report of Public Schools except major capital outlay, debt service, and transportation expenditures. Tuition receipts have been deducted from operating costs because the data are based only on resident pupils.

State school aid defined as "Adjusted state allocation"—this includes the "operating cost allocation" (the basic foundation formula) and "program cost allocation" (a matching program for special education, transportation, vocational, and other programs), plus debt service allocation, and a variety of adjustments.

Pupil count is average resident pupils. Observations are generally districts; but data for 20 community (elementary) districts were incorporated into the regional (high school) districts of which the communities are members.

Massachusetts

Operating spending defined as "Integrated Operating Cost"—direct school spending plus spending outside the school budget that benefits schools such as insurance and pupil support services; also includes EEO grant spending but not other non-general fund expenditures. Takes into account a town's share of regional school district spending as well as its own schools.

State school aid defined as "State revenue"—the sum of Chapter 70, school construction, pupil transportation, state wards, food services, Chapter 188 (EEO grants and other phased-out grants), special education Chapter 71b, racial imbalance, regional school aid (matching inversely to wealth per pupil), and other state aid.

Pupil count is net average membership. Observations are cities and towns. Aid and spending for regional districts, vocational districts, and county agricultural schools were allocated back to member cities and towns on the basis of enrollments.

New Hampshire

Operating spending defined as "K–12 Cost"—current expenditures reported on each district's annual financial report. These represent operating costs and do not include tuition, transportation, capital expenditures, or debt service expenditures. Food service is deducted from current expenditures.

State school aid defined as "Foundation aid"—this excludes building aid, state contributions to teacher pensions, special education catastrophic aid, and vocational education tuition and transportation aid.

Pupil count is average daily membership in residence. Observations are school districts, many of which coincide with cities and towns.

Rhode Island

Operating spending defined as "Expenditures, All Programs"—includes general instruction and special programs such as vocational, special education, limited English-proficient, compensatory, and gifted/talented.

State school aid defined as "State revenue"—includes operations aid (basic equalizing program), disadvantaged program, vocational education, gifted, and other aid.

Pupil count is resident average daily membership. Observations are school districts, which generally coincide with cities and towns. Two regional districts were deleted because of incomplete data.

Vermont

Operating spending defined as "Current Expenditures" —all school expenditures (1) excluding capital debt service expenditures, transportation expenditures, and special education expenditures, and (2) subtracting incoming tuition and all federal and state funds earned except for federal impact aid and state general aid.

State school aid defined as "Calculated general state aid"—the sum of basic aid (foundation formula) plus minimum aid adjustment plus two kinds of supplemental aid (capital debt service and above-average expenditures) plus maximum loss adjustment.

Pupil count is average daily membership. Observations are districts, many of which coincide with cities and towns.

References

Berne, Robert. 1988. "Equity Issues in School Finance." *Journal of Educational Finance*, vol. 14 (Fall), pp. 159–80.

Berne, Robert and Leanna Stiefel. 1984. *The Measurement of Equity in School Finance: Conceptual, Methodological, and Empirical Dimensions*. Baltimore and London: The Johns Hopkins University Press.

Bradbury, Katharine L. 1988. "Shifting Property Tax Burdens in Massachusetts." *New England Economic Review*, September/October, pp. 36–48.

Coons, John E., William H. Clune III, and Stephen D. Sugarman. 1970. *Private Wealth and Public Education*. Cambridge, MA: The Belknap Press of Harvard University Press.

Council of Chief State School Officers (Ramsay Selden and Todd Landfried). 1988. *State Education Indicators 1988*.

Downes, Thomas A. 1992. "Evaluating the Impact of School Finance Reform on the Provision of Public Education: The California Case." *National Tax Journal*, vol. XLV, no. 4 (December), pp. 405–19.

Feldstein, Martin S. 1975. "Wealth Neutrality and Local Choice in Public Education." *American Economic Review*, vol. 65, no. 1 (March), pp. 75–89.

Gold, Steven D., David M. Smith, Stephen B. Lawton, and Andrea C. Hyary (editors). 1992. *Public School Finance Programs of the United States and Canada 1990–91*. American Education Finance Association and Center for the Study of the States, The Nelson A. Rockefeller Institute of Government, State University of New York.

Kosterlitz, Julie. 1990. "Schoolhouse Equality." *National Journal*, July 21, pp. 1768–1773.

Kozol, Jonathan. 1991. *Savage Inequalities: Children in America's Schools*. New York: Crown Books.

LaMorte, Michael W. 1989. "Courts Continue to Address the Wealth Disparity Issue." *Educational Evaluation and Policy Analysis*, vol. 11, no. 1 (Spring), pp. 3–15.

Legislative Research Commission. 1991. *The Kentucky Education Reform Act of 1990: A Citizen's Handbook*. Frankfort, Kentucky.

National Education Association (Peter Veillette). 1987. *SEARCH: Understanding State School Finance Formulas*. National Education Association of the United States.

Odden, Allan, with the assistance of Van Dougherty. 1984. *Education Finance in the States: 1984*. Education Finance Center, Education Commission of the States, Report F84–1. June.

Odden, Allan, C. Kent McGuire, and Grace Belsches-Simmons. 1983. *School Finance Reform in the States: 1983*. Education Finance Center, Education Commission of the States, Report F83–1. January.

Odden, Allan R. and Lawrence O. Picus. 1992. *School Finance: A Policy Perspective*. New York: McGraw-Hill, Inc.

Reischauer, Robert D. and Robert W. Hartman with the assistance of Daniel J. Sullivan. 1973. *Reforming School Finance*. Washington, D.C.: The Brookings Institution.

Robinson, Lauren. 1992. "SJC takes school case: Lawsuit challenges property-tax funding." *The Boston Globe*. December 15, pp. 25, 28.

U.S. Advisory Commission on Intergovernmental Relations. 1990. *The Structure of State Aid to Elementary and Secondary Education*. Report M-175 (December).

U.S. Department of Education, National Center for Education Statistics. 1992. *Digest of Education Statistics 1992*. October.

Verhover, Sam Howe. 1993. "Texas to Hold Referendum on School-Aid Shift to Poor." *The New York Times*, February 16, p. A12.

Walker, Billy D. 1990. "The Impact of School Finance Litigation in Texas, 1968–1990." *Texas Researcher*, vol. 1 (Fall), pp. 1–17.

Wise, Arthur E. 1968. *Rich Schools, Poor Schools: The Promise of Equal Educational Opportunity*. Chicago and London: The University of Chicago Press.

Wong, Doris Sue. 1993. "SJC hears arguments in suit to change education funding." *The Boston Globe*. February 3, p. 20.

Forecasting Investment with Models and Surveys of Capital Spending

The investment spending of businesses is one of the most closely watched elements of the national accounts, partly because variations in this capital spending account for much of the variation in national output during business cycles, and partly because future living standards depend on the volume of previous capital formation. Because accurate forecasts of investment are especially valuable for policymakers, the U.S. Department of Commerce has collected surveys of businesses' plans for capital spending for almost half a century. At their inception, these surveys were essential for forecasting investment. In the 1940s and 1950s, statistical forecasting was a nascent art, and national income accounting was only beginning to provide the data required for statistical modelling. Despite the great strides in statistical forecasting during the ensuing decades, the surveys retain considerable appeal because the magnitude of the errors of the models can be uncomfortably large for policymakers.

This article assesses the recent contribution of the Department of Commerce's survey to forecasts of business investment, once other information on business conditions is taken into account. In principle, the survey ought to be valuable. Economic data might reveal much about prevailing business conditions, but the survey can record investors' intentions. The scope of the survey is limited, however, and the relationship between the capital spending of the survey's respondents and the investment spending reported in the national accounts can vary considerably.

Despite its promise, the survey appears to improve forecasts of investment only marginally since the 1970s. For forecasts as short as one quarter, knowing the survey's results is not as valuable as knowing the history of investment spending and the output of businesses. For forecasts of investment over the coming year, the information in the survey is not as useful as that in the history of output, cash flows, costs of capital, and investment itself.

Richard W. Kopcke

Vice President and Economist, Federal Reserve Bank of Boston. *The author thanks Mark Howrey for his research assistance.*

The survey has failed to fulfill its potential since the 1970s mainly because the capital spending of the businesses that respond to the survey has increased significantly relative to the investment spending reported in the national accounts. In the past, when the ratio of the respondents' capital spending to total investment spending was more stable, the survey was a more accurate indicator of businesses' purchases of structures and equipment. If this ratio should become more stable in the future, then the survey might become a more reliable indicator of the aggregrate investment spending of businesses.

The rate of growth of investment spending cannot be predicted with great precision.

This article's results imply that the rate of growth of investment spending cannot be predicted with great precision. For example, if investment is expected to grow 10 percent from one year to the next, the actual rate of growth could easily be as great as 15 percent or as low as 5 percent. This uncertainty blurs the apparent distinctions among forecasts. It also suggests that policymakers who wish to guarantee a rapid rate of capital formation may need to set their sights very high indeed to be confident of success.

I. The Census Bureau's Surveys of Capital Spending

Forecasts of investment spending by businesses depend not only on projections from statistical models but also on surveys of the capital budgets of businesses, principally that conducted by the U.S. Bureau of the Census. This survey covers expenditures by businesses for new plants and equipment that are to be used in the United States.

These surveys, done quarterly since 1947, collect planned expenditures one, two, and three quarters in advance. In the fourth quarter of each year, the survey also collects planned spending for the ensuing calendar year. The Census Bureau currently collects responses for the quarterly surveys from approximately 5,000 companies; for the annual survey conducted at the end of each year, the Bureau receives responses from approximately 9,300 companies.

The capital spending reported by the survey's respondents does not necessarily correspond to the concept of business investment reported in the national accounts. The Census samples enterprises in manufacturing and in the mining, transportation, public utility, and commercial industries. The quarterly survey does not cover farm enterprises, professional organizations, or real estate operators. The scope of the annual survey is somewhat more comprehensive, covering real estate operators, hospitals, and religious enterprises. The Bureau uses the results of the surveys—adjusted for outliers, seasonal factors, and systematic biases—in conjunction with benchmarks for past investment to estimate planned plant and equipment spending for each quarter.

Comparing the spending for the coming year reported in the Bureau's fourth-quarter survey to the actual spending subsequently reported by these respondents shows that the survey can predict the rate of growth of this nominal capital spending fairly accurately (Table 1). During the 1980s and early 1990s, a period when the models failed to predict accurately the construction of nonresidential structures, this survey often predicted its measure of capital spending more accurately than most of the models predicted total nominal investment spending as reported in the national accounts. The survey misstated the annual growth of capital spending with an average absolute error of approximately 2.5 percentage points; the models misstated the annual growth of investment with average absolute errors exceeding 6 percentage points.

Although the Census survey often predicted its measure of capital spending comparatively accurately, the survey has not necessarily predicted aggregate investment spending nearly as well, because the relationship between the capital spending reported by the survey's respondents and the investment spending reported in the national accounts has been changing, especially since the 1970s. From 1960 to 1980, the actual capital spending reported by the respondents to the Census survey varied between 77 percent and 85 percent of the nonresidential investment reported in the national accounts. Between 1981 and 1991, this ratio increased almost steadily from 80 percent to 96 percent.

Results such as those shown in Table 1 do suggest that the Census survey, nonetheless, may be an important ingredient for forecasting accurately the investment spending of businesses. The demand for

Table 1
Forecast Errors for Projections of Nominal Investment Spending[a]

	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Survey Data:												
Census Bureau	−2.3	−1.2	−8.8	−.6	−1.5	.5	−5.0	−1.0	1.5	2.2	−2.7	−3.4
Models:												
Accelerator	3.6	10.0	13.8	10.8	12.4	11.6	2.9	−1.9	−2.9	−6.3	−6.6	−9.4
Neoclassical	8.5	15.4	17.2	11.8	10.2	6.5	−2.5	−4.6	−4.7	−7.1	−4.3	−4.7
q Model	8.0	7.7	1.6	−2.2	6.6	9.4	−.1	−6.8	−6.2	−3.5	−8.1	−16.5
Cash Flow	8.5	13.4	13.9	6.6	9.2	10.3	9.2	2.2	−1.6	2.7	10.9	11.3
Autoregression												
with output	9.1	14.8	15.8	8.2	8.2	7.9	1.5	−.6	.8	−.1	2.4	1.9
without output	7.2	12.0	8.2	5.0	18.5	23.5	20.1	19.7	24.7	26.3	27.2	21.9

[a]The nominal Census Bureau annual survey data are compared to the reported nominal investment by the respondents to the survey. The model forecasts are the average of the quarterly model forecasts for equipment and nonresidential structures multiplied by their respective deflators found in Kopcke (1993). The errors shown are actual less predicted as a percentage of actual investment.
Source: Survey data provided by the U.S. Bureau of the Census.

investment goods, after all, depends on investors' expectations of the future. Although investors' views are grounded in their previous experiences, encouraging forecasters to consult previous sales, profits, and costs in order to predict future capital spending, these predictions might benefit considerably by incorporating the direct measures of investors' sentiments that are reported in the Census survey.

Even if surveys of businesses' plans for capital spending provided the most accurate forecasts of investment spending, statistical models would remain useful analytical tools. Statistical models may forecast investment spending over intervals of time not covered by surveys. Moreover, the models may describe the influence of economic conditions on investment spending, so that forecasts may adapt when these conditions change. Finally, some models permit policymakers to assess the potential consequences of changing monetary or fiscal policies. The collection of survey data rich enough to satisfy these objectives, even if feasible, is impractical.

II. The Contribution of Surveys to the Forecasts of Models of Investment Spending

Surveys and statistical models often complement one another. Few forecasts, for example, rest on the projections of models alone. Often the projections of statistical models are adjusted according to forecasters' assessments of business conditions. These adjustments may reflect the results of the Census surveys.

The potential value of surveys may not be the same for all statistical models of investment. If the sentiment recorded in a survey were correlated closely with previous changes in output, for example, then those models that use past output to forecast investment might benefit less from the survey than other models. Furthermore, because the models for purchases of nonresidential structures have performed relatively poorly since the mid 1970s, these models might gain more from surveys than models for purchases of producers' durable equipment.

This section examines the potential contribution of surveys to the five statistical models of investment (Table 2) presented in a previous article, "The Determinants of Business Investment" (Kopcke 1993). For this article, the models were modified so that they consider only previous values of their explanatory variables—output, cash flow, and the cost of capital. Their forecasts of investment this quarter do not depend on the values of their explanatory variables this quarter.

Each model of investment spending was estimated three times (see the Appendix) using data from 1960 to the late 1970s. The first version includes only lagged values of output, cash flow, or the cost of capital as explanatory variables for the accelerator, neoclassical, q, or cash flow models. For the autoregression, lagged values of investment were included, and then both lagged investment and lagged output. The second version is identical to the first except that each model also includes the anticipated capital spending as reported in the Census survey. The third

Table 2
The Models of Investment

Accelerator

$$I_t = a + \sum_{i=0}^{n} b_i Q_{t-i} + cK_{t-1}$$

Neoclassical

$$I_t = a + \sum_{i=0}^{n} b_i \frac{Q_{t-i}}{UCC_{t-i-1}} + \sum_{i=0}^{n} c_i \frac{Q_{t-i}}{UCC_{t-i}} + dK_{t-1}$$

q

$$I_t = a + \sum_{i=0}^{n} b_i[(q-1)_{t-i}K_{t-1-i}] + cK_{t-1}$$

Cash Flow

$$I_t = a + \sum_{i=0}^{n} b_i \left(\frac{F}{C}\right)_{t-i}$$

Autoregression

$$I_t = a + \sum_{i-1}^{n} b_i I_{t-i}$$

Explanation of Symbols
C: price index for capital goods
F: cash flow
I: real investment
K: real stock of capital
Q: real output
q: ratio of financial market valuation of assets to the replacement cost of assets
UCC: user cost of capital

version resembles the second except that the previous period's investment spending replaces the spending reported in the Census survey.

These three versions of each model are used to forecast investment spending from the late 1970s to the present. Comparing the models' projections of investment spending helps isolate the potential contribution of the Census survey data to forecasts of investment. The forecasts from the second version of each model should be more accurate than those from the first version if the survey includes information that is not already represented in the models' other explanatory variables. This comparison, however, may overstate the contribution of the survey for at least two reasons.

First, the omission of contemporaneous output, cash flow, and capital costs may penalize the first version of each model too greatly. Although the actual values of these variables may not be known at the time of the forecast, reasonably accurate projections may be available. This potential penalty does not appear to be substantial for the quarterly forecasts during the 1980s and early 1990s, but it may be more important for the annual forecasts.[1]

Second, the survey data may reduce forecast errors simply because they prevent the models from wandering off course. The surveys reflect the recent amount of capital spending—this period's anticipated spending very much resembles last period's outlays. No model, other than the autoregression, includes investment spending among its explanatory variables; consequently, when these models stray off course, the survey, as a surrogate for recent investment spending, allows them to make midcourse corrections. In this case, the survey makes no unique contribution, because this midcourse correction might be accomplished in other ways—for example, by including the investment spending from the previous quarter instead of the survey among the explanatory variables. If the survey data contain unique, timely information, then the second version of each model should tend to be more accurate than the third version (which replaces the survey data with lagged investment). If the surveys essentially are surrogates for the recent amount of investment spending, then the third version of each model may be most accurate.

III. Quarterly Forecasts

The single most important ingredient for accurate one-quarter forecasts of investment spending during the 1980s and early 1990s was the value of investment spending during the preceding quarters. The value of adding surveys of planned capital spending to models that already included output, cash flow, or the cost of capital was not as great as the value of adding the previous quarter's investment spending. Furthermore, once prior investment had been taken into account, other variables—output, cash flows, and the cost of capital from previous quarters—tended to contribute comparatively little to the accuracy of forecasts.

[1] The quarterly forecasts presented in Kopcke (1993) take full advantage of the contemporaneous values of these explanatory variables. The quarterly forecasts described below do not. Yet the errors of these two sets of forecasts are very similar, according to the statistics reported in Table 7 of the previous article and in the uppermost panels of Tables 4 and 5 of this article.

Table 3
Selected Statistics of the Models for Quarterly Investment in Equipment, for the Estimation Period 1962:I to 1979:IV

Model	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion	Autocorrelation Coefficient	Number of Lags
			Without Survey Data			
Accelerator	5.8	7.2	27.8	6.9	.61	3
Neoclassical	4.1	5.0	12.9	.0	.42	13
q Model	11.8	14.0	62.5	44.4	.86	5
Cash Flow	9.9	11.8	52.8	25.0	.70	5
Autoregression						
without output	3.5	4.7	11.1	1.4	n.a	4
with output	3.1	4.1	4.2	1.4	n.a.	4
			With Survey Data			
Accelerator	5.3	6.6	19.4	4.2	.47	3
Neoclassical	3.7	4.4	2.9	.0	.28	13
q Model	8.1	10.1	45.8	22.2	.59	5
Cash Flow	6.4	8.5	29.2	11.1	.39	5
Autoregression						
without output	3.5	4.7	12.5	2.8	n.a.	4
with output	3.1	4.1	4.2	.0	n.a.	4
			With Lagged Investment in Equipment			
Accelerator	3.2	4.4	5.6	1.4	.00	3
Neoclassical	3.2	3.9	4.3	.0	.01	13
q Model	3.3	4.8	11.1	1.4	.05	5
Cash Flow	3.6	5.0	11.1	5.6	.02	5

Tables 3 and 4 describe the models' errors for purchases of producers' durable equipment; Tables 5 and 6 describe the errors for purchases of nonresidential structures. The first table in each pair summarizes the models' errors during the period of estimation from the early 1960s to the late 1970s. The second table summarizes the models' forecast errors from the late 1970s to the present. Figure 1 shows the correspondence between the models' descriptions of spending on equipment and actual outlays; Figure 2 compares the descriptions of spending on structures with actual outlays.

For explaining purchases of equipment, the accelerator, neoclassical, and autoregression models generally benefited negligibly from the survey. During the estimation period, the mean absolute errors and root mean squared errors fell only slightly, at most, for these four models after introducing the survey variable (Table 3, uppermost and middle panels). By most measures, the accelerator model tended to forecast equipment spending more accurately when the survey was *not* included, while the performance of the neoclassical and autoregression models improved only slightly by including the survey. Nevertheless, the surveys might have been valuable at times. Since 1989, for example, the accelerator model's forecasts of purchases of equipment benefited from the survey (Figure 1).

The cash flow model and the q model, which do not include a measure of output, benefited considerably more from including the survey. During the estimation period, the mean absolute errors and root mean squared errors for the cash flow and q models were reduced by almost one-third after including survey information (Table 3). During the forecast period, the survey reduced these error statistics by approximately one-half for the cash flow model (Table 4). Since 1986, the surveys made the forecasts of the q model and the cash flow model for equipment considerably more accurate (Figure 1).

Table 4
Selected Statistics of the Models for Quarterly Forecasts of the Investment in Equipment, 1980:I to 1992:I

Model	Mean Error	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion
			Without Survey Data		
Accelerator	−8.9	13.8	16.6	67.3	49.0
Neoclassical	5.2	10.9	13.4	61.2	36.7
q Model	−22.2	23.2	27.3	85.7	69.4
Cash Flow	16.2	19.2	26.5	73.5	42.9
Autoregression					
without output	1.1	6.0	7.7	30.6	8.2
with output	−2.0	5.5	6.9	18.4	6.1
			With Survey Data		
Accelerator	.5	17.0	18.5	87.8	73.5
Neoclassical	−5.2	9.3	11.8	49.0	30.6
q Model	−15.3	20.2	24.4	73.5	67.3
Cash Flow	−5.6	9.8	12.0	55.1	24.5
Autoregression					
without output	.4	6.0	7.6	26.5	8.2
with output	−1.3	5.3	6.7	18.4	4.1
			With Lagged Investment in Equipment		
Accelerator	−2.8	5.7	7.6	24.5	10.2
Neoclassical	.5	5.7	7.3	32.7	6.1
q Model	−7.7	9.1	11.3	51.0	24.5
Cash Flow	.2	5.4	7.0	22.4	4.1

For structures, the surveys reduced the errors of all models, but even with the assistance of the surveys, the performance of all the models, except the autoregression, remained very poor. The errors of all the models, with or without the survey data, generally were small during the estimation period (Table 5 and Figure 2). During the forecast period, however, the errors of the models often were large, especially in the late 1970s and early 1980s (Table 6 and Figure 2). Although the inclusion of the surveys reduced these errors substantially for all models except the neoclassical and autoregression, the errors remained very great. The surveys reduced the models' average forecast errors during the early 1980s, but by the early 1990s the surveys generally increased the models' errors in forecasting investment in structures.

For these one-quarter forecasts, the models benefited more from the inclusion of lagged investment spending than they did from the inclusion of the surveys. For both equipment and structures, the average forecast errors of all models including lagged investment spending were only a fraction of the average errors for the models that included the survey data (Tables 4 and 6). Furthermore, for both equipment and structures, the profile of these forecasts generally resembled that of investment spending (Figures 1 and 2).[2]

Previous investment not only contributed more to the accuracy of one-quarter forecasts than did the surveys, previous investment also contributed more than lagged cash flow or the cost of capital. The error statistics for the autoregression (Tables 4 and 6, uppermost panels) are generally at least as low as those for the versions of the other models that in-

[2] By including lagged *total* investment in the equations for equipment rather than the survey data, the average errors for most models generally were only marginally lower than those in the second panel of Table 4; however, the average errors for the cash flow equation were approximately one-fifth greater. By including *total* investment rather than the survey data in the equations for structures, the average errors were substantially less than those in the second panel of Table 5, but not as low as those in the lowest panel.

Figure 1

Actual Expenditures and Models' One-Quarter Forecasts of Investment in Equipment

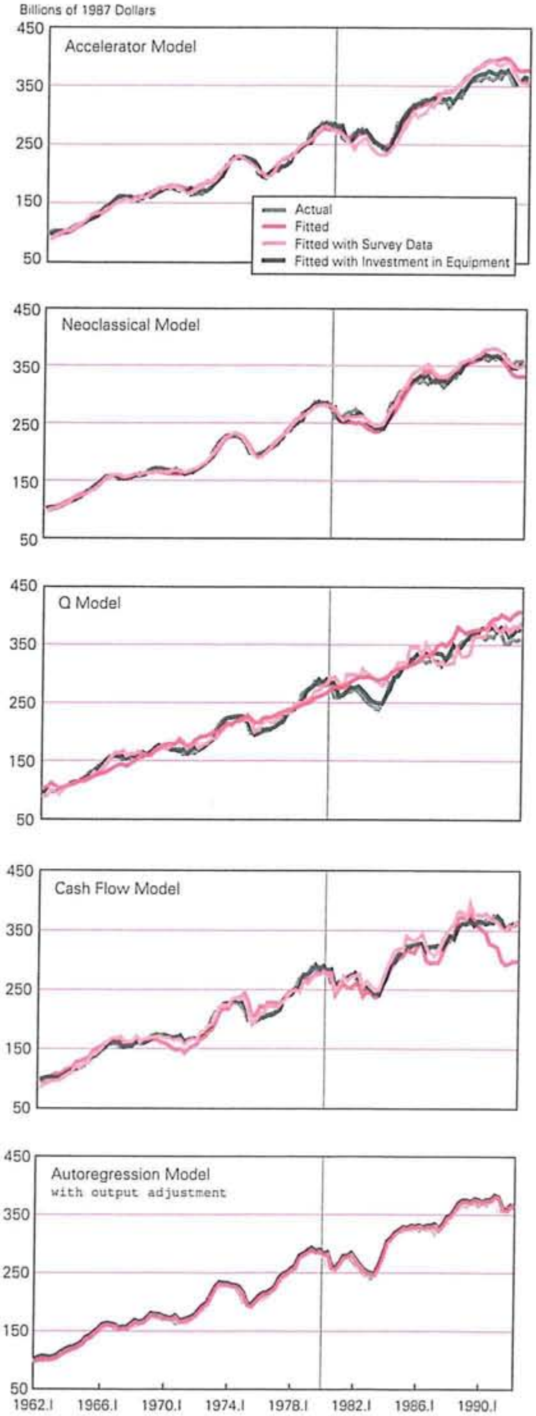


Figure 2

Actual Expenditures and Models' One-Quarter Forecasts of Investment in Structures

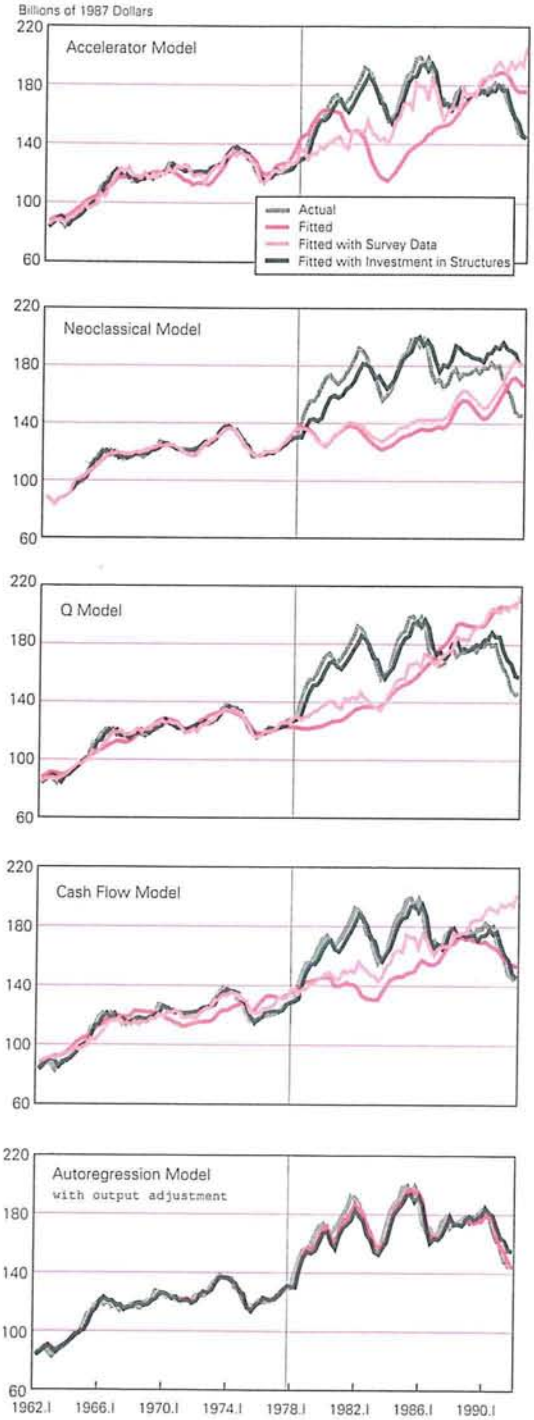


Table 5

Selected Statistics of the Models for Quarterly Investment in Nonresidential Structures, for the Estimation Period 1962:I to 1977:IV

Model	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion	Autocorrelation Coefficient	Number of Lags
			Without Survey Data			
Accelerator	4.8	5.6	10.9	1.6	.81	11
Neoclassical	2.2	2.7	.0	.0	.41	20
q Model	3.3	4.1	4.7	.0	.73	8
Cash Flow	6.5	7.4	31.3	6.3	.89	12
Autoregression						
without output	1.9	2.4	.0	.0	n.a.	4
with output	1.6	2.2	.0	.0	n.a.	4
			With Survey Data			
Accelerator	2.8	3.5	3.1	.0	.35	11
Neoclassical	2.2	2.7	.0	.0	.44	20
q Model	2.1	2.8	3.1	.0	.51	8
Cash Flow	3.9	4.7	7.8	.0	.76	12
Autoregression						
without output	1.9	2.3	.0	.0	n.a.	4
with output	1.6	2.1	.0	.0	n.a.	4
			With Lagged Investment in Structures			
Accelerator	1.7	2.2	.0	.0	.00	11
Neoclassical	1.4	1.7	.0	.0	.04	20
q Model	1.8	2.2	.0	.0	.01	8
Cash Flow	1.8	2.4	.0	.0	.02	12

Table 6

Selected Statistics of the Models for Quarterly Forecasts of the Investment in Nonresidential Structures, 1978:I to 1992:I

Model	Mean Error	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion
			Without Survey Data		
Accelerator	16.9	25.0	31.5	75.4	59.6
Neoclassical	32.1	35.1	38.5	94.7	89.5
q Model	11.7	31.3	35.1	96.5	86.0
Cash Flow	20.8	21.8	27.2	73.7	56.1
Autoregression					
without output	2.4	4.5	5.4	14.8	1.6
with output	2.0	4.5	5.4	16.4	1.6
			With Survey Data		
Accelerator	8.4	19.7	23.7	77.2	66.7
Neoclassical	26.8	31.9	35.0	94.7	87.7
q Model	8.6	25.7	29.9	84.2	78.9
Cash Flow	7.8	18.8	22.8	75.4	61.4
Autoregression					
without output	2.5	5.3	6.5	32.8	.0
with output	3.0	5.9	7.2	45.9	26.2
			With Lagged Investment in Structures		
Accelerator	2.6	4.7	5.7	19.3	1.8
Neoclassical	−3.6	12.5	15.3	66.7	36.8
q Model	.3	6.7	8.0	42.1	5.3
Cash Flow	2.2	4.7	5.7	15.8	.0

clude the previous quarter's investment spending (lowest panels). For predicting purchases of equipment, the version of the autoregression that includes lagged output was more accurate, but for predicting purchases of structures, adding output to the autoregression alters the error statistics negligibly.

IV. Annual Forecasts

Whereas transient influences may account for much of the quarterly change in investment spending, the role of more fundamental determinants of investment becomes more important over longer forecast horizons. For example, investment is not likely to diverge very much from its customary share of GDP for very long unless either the return on capital rises considerably or the cost of capital falls significantly. For the models that rely on output, cash flow, or the cost of capital to predict investment, the addition of lagged investment reduced the average errors of the one-year forecasts much less than it reduced those for the one-quarter forecasts. For purchases of equipment, the accuracy of the models generally improved negligibly, at best, after the inclusion of lagged investment.

Because these tests constrain all models to rely on lagged values of their explanatory variables, these results may overstate the contribution of surveys and lagged investment in these models. As the forecast horizon expands to a year or more, the course of investment in the future often deviates considerably from that implied by previous investment. Accordingly, informed judgments about prospective GDP, profits, and the cost of capital may become essential ingredients for longer-run forecasts.[3]

Tables 7 and 8 describe the models' errors for purchases of producers' durable equipment; Tables 9 and 10 describe the errors for purchases of nonresidential structures. The first table in each pair summarizes the models' errors during the period of estimation ending in the late 1970s. The second table summarizes the models' forecast errors from the late 1970s to the present. Figure 3 shows the correspondence between the models' descriptions of spending on equipment and actual outlays; Figure 4 compares the descriptions of spending on structures with actual outlays.

For equipment, the accelerator, neoclassical, and autoregression models benefited little from the inclusion of the surveys, while the q and cash flow models conformed to the course of investment more closely after taking the surveys into account (Figure 3). The cash flow model benefited the most from the survey (Tables 7 and 8, uppermost and middle panels). During the estimation period its mean absolute error and its root mean squared error fell by one-quarter after adding the survey. During the forecast period, these average errors fell by one-third.

For structures, all models except the autoregression benefited to a degree by taking the surveys into account. During the estimation period, the surveys reduced the error statistics most for the accelerator and cash flow models (Table 9). During the forecast period, the average errors of the q and cash flow models fell the most after the inclusion of the surveys (Table 10). The autoregression's average error in forecasting investment in structures increased substantially after the survey was added to its equations.

For the one-year forecasts, the versions of the models using the previous year's investment instead of the surveys often were more accurate than the versions that used the surveys. But, for these one-year forecasts, the models generally benefited less from this substitution than they did for the one-quarter forecasts. The average errors for equipment were nearly identical for the survey and the lagged

The role of fundamental determinants of investment becomes more important over longer forecast horizons.

investment versions of the models (Table 8, middle and lowest panels).

Although the average errors for the forecasts of structures were lower when lagged investment replaced the surveys (Table 10), the forecasts that included the surveys often anticipated more accurately the turning points in the course of this investment (Figure 4). The significance of this observation is qualified, however. By the design of these tests, the forecasts that do not include the surveys use the average values of variables dated a year or more

[3] The extended forecasts presented in Kopcke (1993) show that the projections from the simple autoregression (without an adjustment for output) stray from the course of investment and that the profile of these projections does not resemble very closely that of actual investment.

Table 7

Selected Statistics of the Models for Annual Investment in Equipment, for the Estimation Period 1962 to 1979

Model	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion	Autocorrelation Coefficient	Number of Lags
			Without Survey Data			
Accelerator	8.4	10.0	44.4	22.2	.18	0
Neoclassical	3.5	4.2	6.3	.0	.02	4
q Model	11.8	14.5	66.7	50.0	.29	0
Cash Flow	11.8	15.7	50.0	38.9	.16	2
Autoregression						
without output	8.9	9.4	61.1	11.1	n.a.	3
with output	5.1	5.8	16.7	.0	n.a.	3
			With Survey Data			
Accelerator	8.1	9.9	44.4	22.2	.13	0
Neoclassical	3.5	4.2	6.3	.0	.02	4
q Model	8.1	10.2	38.9	27.8	.10	0
Cash Flow	9.5	11.5	50.0	27.8	.15	2
Autoregression						
without output	8.6	9.2	61.1	5.6	n.a.	3
with output	5.1	5.7	16.7	.0	n.a.	3
			With Lagged Investment in Equipment			
Accelerator	7.9	9.5	44.4	27.8	.27	0
Neoclassical	2.5	3.7	6.3	.0	.0	4
q Model	10.0	12.9	55.6	33.3	.17	0
Cash Flow	9.7	13.4	38.9	33.3	.03	2

Table 8

Selected Statistics of the Models for Annual Forecasts of the Investment in Equipment, 1980 to 1991

Model	Mean Error	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion
			Without Survey Data		
Accelerator	−10.1	15.7	19.3	58.3	50.0
Neoclassical	2.1	12.3	14.2	75.0	50.0
q Model	−26.2	26.2	30.5	83.3	75.0
Cash Flow	14.0	20.4	24.3	83.3	66.7
Autoregression					
without output	−6.8	15.7	20.3	75.0	41.7
with output	−7.4	14.6	16.1	83.3	58.3
			With Survey Data		
Accelerator	−6.1	16.2	19.2	66.7	58.3
Neoclassical	1.0	12.3	14.1	75.0	50.0
q Model	−12.9	22.3	26.3	91.7	66.7
Cash Flow	−8.5	13.6	16.1	75.0	58.3
Autoregression					
without output	−10.0	15.0	20.9	50.0	41.7
with output	−6.5	14.3	15.9	75.0	58.3
			With Lagged Investment in Equipment		
Accelerator	−10.1	16.3	20.2	66.7	50.0
Neoclassical	7.6	13.1	16.0	75.0	25.0
q Model	−19.1	21.3	24.4	91.7	83.3
Cash Flow	.5	12.7	15.4	75.0	25.0

Figure 3

Actual Expenditures and Models' One-Year Forecasts of Investment of Equipment

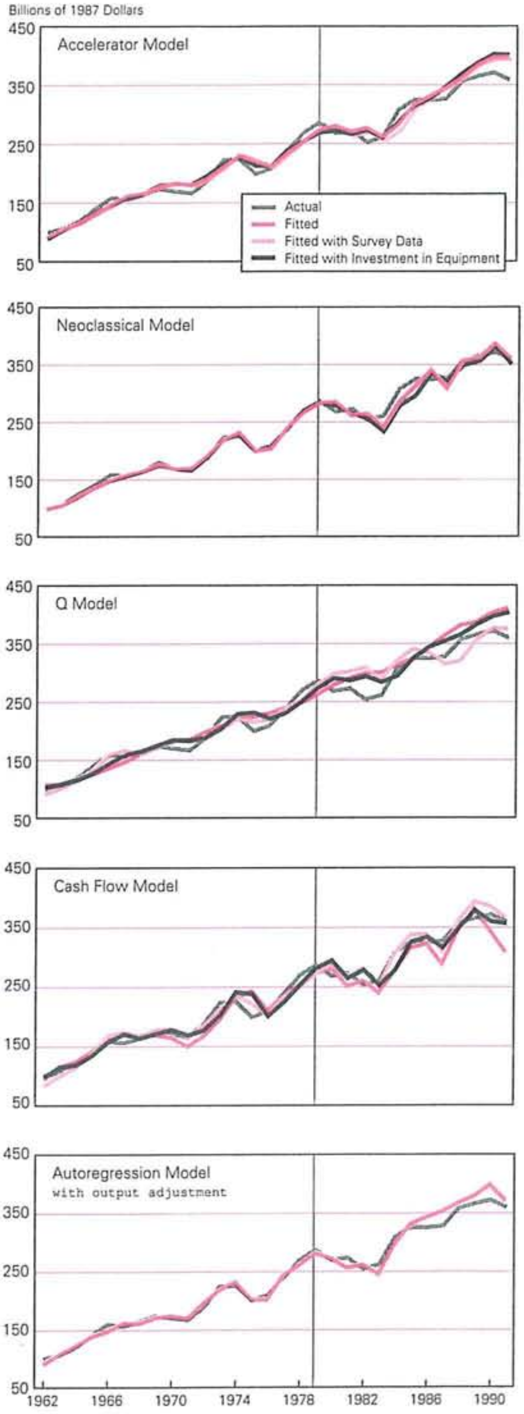


Figure 4

Actual Expenditures and Models' One-Year Forecasts of Investment in Structures

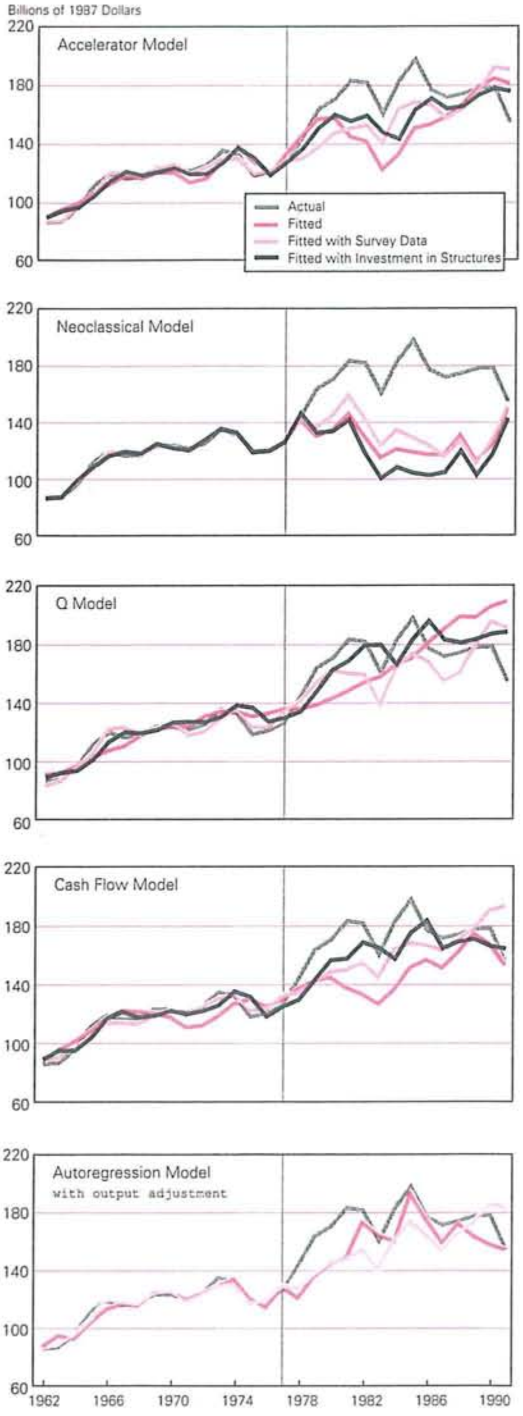


Table 9

Selected Statistics of the Models for Annual Investment in Nonresidential Structures, for the Estimation Period 1962 to 1977

Model	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion	Autocorrelation Coefficient	Number of Lags
			Without Survey Data			
Accelerator	5.2	5.9	18.8	.0	.16	0
Neoclassical	1.5	1.9	.0	.0	.08	4
q Model	5.4	6.9	31.3	6.3	.50	0
Cash Flow	6.6	7.7	31.3	6.3	.34	2
Autoregression						
without output	3.7	4.4	6.3	.0	n.a.	3
with output	3.2	4.0	6.3	.0	n.a.	3
			With Survey Data			
Accelerator	2.1	2.6	.0	.0	.01	0
Neoclassical	1.2	1.7	.0	.0	.12	4
q Model	3.5	4.3	6.3	.0	.07	0
Cash Flow	3.3	3.9	.0	.0	.27	2
Autoregression						
without output	3.0	3.6	.0	.0	n.a.	3
with output	1.9	2.5	.0	.0	n.a.	3
			With Lagged Investment in Structures			
Accelerator	4.2	5.2	12.5	.0	.04	0
Neoclassical	1.5	1.9	.0	.0	.05	4
q Model	5.1	6.5	12.5	6.3	.15	0
Cash Flow	4.2	5.5	18.8	6.3	.01	2

Table 10

Selected Statistics of the Models for Annual Forecasts of the Investment in Nonresidential Structures, 1978 to 1991

Model	Mean Error	Mean Absolute Error	Root Mean Squared Error	Percent of Absolute Errors Exceeding $8 Billion	Percent of Absolute Errors Exceeding $13 Billion
			Without Survey Data		
Accelerator	16.9	22.2	27.8	64.3	50.0
Neoclassical	44.3	44.3	48.9	85.7	85.7
q Model	.9	22.9	26.3	78.6	78.6
Cash Flow	23.5	23.5	28.6	78.6	64.3
Autoregression					
without output	10.9	13.2	16.4	71.4	42.9
with output	13.0	13.8	17.4	57.1	50.0
			With Survey Data		
Accelerator	12.5	19.8	22.0	92.9	78.6
Neoclassical	38.6	38.6	43.5	85.7	85.7
q Model	7.7	15.9	18.4	78.6	64.3
Cash Flow	9.8	17.5	20.6	78.6	57.1
Autoregression					
without output	9.5	18.0	20.8	71.4	71.4
with output	13.5	19.1	21.0	85.7	71.4
			With Lagged Investment in Structures		
Accelerator	12.6	15.6	19.3	71.4	35.7
Neoclassical	53.2	53.4	59.0	92.9	85.7
q Model	−1.4	13.4	15.3	78.6	50.0
Cash Flow	10.3	13.4	15.2	64.3	42.9

before the date of the forecast. Therefore, these forecasts cannot benefit from the more current information that often is necessary for timing turning points accurately. The forecasts that include the annual survey by the Census Bureau, however, incorporate more timely information, because these surveys are taken just before the beginning of the year. Accordingly, these forecasts should tend to anticipate turning points more accurately.

The one-year results suggest that accurate longer-term forecasts might depend on variables other than previous investment spending and output. For the one-quarter forecasts, the average errors for the autoregression were only a fraction of those for the other models, but for the one-year forecasts of purchases of equipment, the average errors of the autoregressions were more nearly comparable to those of the other models (Table 8, uppermost panel).[4] Adding either the surveys or the previous year's purchases of equipment to these models altered the average forecast errors only negligibly for the accelerator and neoclassical models, while the average errors for the q and cash flow models were reduced by one-fifth and one-third, respectively. The autoregression predicted annual purchases of structures more accurately than other models (Table 10, uppermost panel); however, compared to the results for the one-quarter forecasts, the autoregression has lost much of its relative advantage.

V. Conclusions

Surveys of plans for capital spending, in principle, are a promising ingredient for forecasts of investment. Statistical models of investment, as helpful as they are for both projecting and analyzing the flow of investment, nevertheless produce uncomfortably large errors, in the opinion of many policymakers. If we were confident that purchases of producers' durable equipment in 1993 will be 15 percent greater than they were in 1992, the need for an investment tax credit might not seem so compelling. The prospect of a 5 percent increase in equipment spending, however, might foster considerable interest in tax incentives for investors. Inasmuch as the magnitudes of average annual forecast errors for statistical models are so substantial for purchases of equipment and for purchases of nonresidential structures, these models too frequently cannot reassure policymakers that prospective investment spending will meet their goals.

Despite the promise of the Census Bureau's survey of capital spending, its ability to predict investment, as reported in the national accounts, is disappointing. The survey does not cover all types of business or all industries, and the capital spending reported by participating businesses does not necessarily match the concept of investment that is reported in the national accounts. Though this survey often anticipates fairly accurately the actual capital spending reported by its respondents, the information in the survey does not improve the performance of statistical models of aggregate business investment spending, because the relationship between respondents' capital spending and aggregate investment can change considerably from year to year. Since the 1970s, for example, the capital spending covered by the survey has increased significantly relative to aggregate investment.

When conditions are especially unsettled, forecasters require statistical models to form a consistent forecast.

For forecasts as short as one quarter, the information in the Census survey has not been as valuable as that inherent in the data for investment spending during previous quarters. Quarter-to-quarter changes in investment seem to be dominated by transient influences that are difficult to describe. The statistical models themselves find that knowledge of cash flow, the cost of capital, and perhaps even output during previous quarters may contribute comparatively little to the forecast once previous investment is taken into account.

For forecasts extending over horizons as long as a year or more, the information in the Census survey has not been as valuable as that inherent in a variety of economic data. Over longer horizons, fundamental trends tend to dominate the course of investment

[4] In Kopcke (1993), in forecasts extending over horizons longer than one year, autoregressions were not as accurate as other descriptions of investment that included output, cash flow, or the cost of capital. These other descriptions benefited, of course, from the information contained in the actual values of these other variables. But the results suggest that with reasonably accurate forecasts of output, cash flow, and the cost of capital these models would still possess an edge over autoregressions.

spending. To a degree, knowing the past course of investment, the "inertia" in capital spending, helps predict future investment. But in this case, the role of output, cash flow, and the cost of capital in determining capital spending also becomes more important. The responses recorded in the Census survey reflect less accurately the amount of investment that will be reported in the national accounts one year in advance than they reflect this spending one quarter in advance.

Perhaps, in the future, the relationship between the capital spending covered by the Census survey and the aggregate investment spending of businesses might become more stable, making the survey a more reliable indicator of aggregate investment spending. Even so, policymakers would continue to rely on statistical models for much of their analysis. A more extensive survey might confidently forewarn policymakers that investment will not meet their standards, but it cannot describe the motives of investors, suggesting how different government policies may alter these motives. Furthermore, surveys are fallible. The forecasts of investment revealed in surveys are no better than the respondents' various readings of economic conditions at the time the survey is taken. When conditions are especially unsettled and the readings of respondents are especially discordant, forecasters require statistical models to form a consistent forecast, anticipating how some businesses may revise their outlook for business conditions and alter their plans for capital spending.

References

Commerce Clearing House, Inc. Various years. *U.S. Master Tax Guide.*

Kiviet, Jan F. and Walter Krämer. 1992. "Bias of s^2 in the Linear Regression Model with Correlated Errors." *The Review of Economics and Statistics*, vol. 74, no. 2 (May), pp. 362–65.

Kopcke, Richard W. 1993. "The Determinants of Business Investment: Has Capital Spending Been Surprisingly Low?" *New England Economic Review*, January/ February, pp. 3–31.

New York Stock Exchange. Various years. *Fact Book.*

Seskin, Eugene P. and David F. Sullivan. 1985. "Revised Estimates of New Plant and Equipment Expenditures in the United States, 1947–83." *Survey of Current Business*, vol. 65, no. 2 (February), pp. 21–23.

Yule, G. Udny. 1926. "Why Do We Sometimes Get Nonsense-Correlations Between Time-Series?" *Journal of the Royal Statistical Society*, vol. 89, part I (January), pp.1–64.

Appendix

Sources of Data: All data are from the U.S. Department of Commerce, Bureau of Economic Analysis, National Income and Product Accounts (NIPA) unless otherwise noted. Measures of stocks of assets and flows of goods or services are expressed in 1987 dollars.

IS, IE: Investment in nonresidential structures, and investment in producers' durable equipment, respectively, for all private businesses. The quarterly investment data are expressed at an annual rate.

KS, KE: Capital stock of structures, and equipment, respectively. Quarterly estimates of the stock of capital were derived from year-end stocks by a nonlinear interpolation assuming the perpetual inventory method and assuming a constant quarterly rate of depreciation throughout the year that is consistent with published data for the end of each year.

EXPITOTL, YRAHEAD: Total expected investment for the next quarter and the next year, respectively. Anticipated total investment is taken from U.S. Bureau of the Census, *Plant and Equipment Expenditure and Plans.* Anticipated investment is converted to constant dollars using the implicit price deflator for fixed nonresidential investment.

RGDPBUS: Real gross domestic product for businesses; quarterly data expressed at an annual rate.

F: Cash flow for businesses, using data from the Board of Governors of the Federal Reserve System, Flow of Funds Section, for the nonfinancial corporate business sector. Cash flow is defined as profits less taxes and dividends, with capital consumption adjustment and depreciation allowances plus capital consumption allowances.

CS, CE, CT: Implicit price deflators for nonresidential structures, producers' durable equipment, and total investment.

NYSEBOND: Market value as a percent of par value for all New York Stock Exchange listed bonds. Annual data come from the NYSE *Fact Book* for various years. Quarterly data were derived using a nonlinear interpolation based on the pattern of new Aa utility bond yields.

q: The ratio of the market value of nonfinancial corporations to the replacement value of their net assets. Market value equals equity less farm net worth plus net interest-bearing debt, which is the sum of bank loans, commercial paper, acceptances, finance company loans, U.S. government loans, and adjusted bonds (AB).

$$AB = .5 * MTG + NYSEBOND * (.5 * MTG + TEB + CB)$$

MTG = commercial mortgages

TEB = tax exempt bonds

CB = corporate bonds

The replacement value of net nonfinancial corporate assets equals total assets less profit taxes payable, trade debt, and foreign direct investment in the United States. Except for NYSEBOND, all data come from the Board of Governors of the Federal Reserve System, Flow of Funds.

INFLATN: Rate of inflation expected over the coming five years. For 1980:IV–1992:I, INFLATN is the average of monthly surveys done by Richard Hoey, available from the Board of Governors of the Federal Reserve System, FAME Database. Hoey's survey data were regressed on lagged values of the annual rate of change of the CPI for 1980:IV–1992:I; this equation was used to obtain expectations for the period 1959:I–1980:III.

RE, RS: User cost of capital for equipment, and nonresidential structures.

$$RE = (CE/CT)(.15 + D)(1 - ITC - TAX * WE - .3 * (1 - DEBTE))/(1 - TAX)$$

$$RS = (CS/CT)(.05 + D)(1 - TAX * WS - .3 * (1 - DEBTS))/(1 - TAX)$$

The rate of depreciation is 0.15 for equipment and 0.05 for structures.

D, the discount rate for corporate profits after corporate income taxes, equals the Standard & Poor's dividend/price ratio for common stocks plus an estimate of the real rate of growth of nonfinancial corporate enterprises, a constant 4 percent. This definition of D is inspired by the Gordon growth model for valuing equities.

ITC, the investment tax credit for equipment, and TAX, the statutory effective tax rate paid by U.S. corporations, are taken from the DRI Model of the U.S. Economy. ITC is the weighted average of investment tax credits for autos, office equipment, and other equipment.

WE is the present value of depreciation allowances for equipment using the most "accelerated" formula permitted by law. From 1959:I through 1981:II, equipment was depreciated using Sum of the Year's Digits; from 1981:III through 1986:IV, equipment was depreciated using the Accelerated Cost Recovery System; from 1987:I through 1992:I, equipment was depreciated using the Modified Accelerated Cost Recovery System. Tax life for equipment is the weighted average of the tax lives for different classes of equipment taken from the DRI Model of the U.S. Economy. The nominal discount rate used equals INFLATN times (1 + 0.015); 0.015 represents the assumed real rate of discount (after taxes).

WS is similarly defined for structures. Structures were depreciated according to Sum of the Year's Digits from 1959:I through 1969:II; from 1969:III through 1981:II, structures were depreciated according to the 150 percent Declining Balance Method; from 1981:III through 1986:IV, buildings were depreciated according to the Accelerated Cost Recovery System, and the Modified Accelerated Cost Recovery System was used thereafter. The discount rate used equals INFLATN times (1 + 0.015).

DEBTE and DEBTS are the present value of debt service charges after taxes per dollar borrowed, for equipment and for structures. The interest rate on debt equals the prevailing Aa new utility rate. The maturity of the loan equals the tax life of the capital good. The discount rate is the same as that for WE and WS. DEBT equals unity when the Aa utility rate, after taxes, equals the discount rate; DEBT exceeds unity when the after-tax interest rate exceeds the discount rate.

Annual data are averages of the quarterly data for each year.

Unless otherwise noted, all quarterly regressions for equipment were run from 1962:I to 1979:IV, while those for structures were run from 1962:I to 1977:IV. Similarly, annual regressions for equipment were run from 1962 to 1979, while those for structures were run from 1962 to 1977, unless noted otherwise. Lag coefficients in all models other than the autoregression were constrained to a third-degree polynomial when the lags were sufficiently long. The last lag coefficients for the quarterly cash flow structures equation were constrained to equal zero. Otherwise, the lag coefficients were not constrained.

The regressions were estimated by ordinary least squares with no allowance for autocorrelation of the errors. Inasmuch as the estimated first-order autocorrelation coefficients for the residuals from these equations could be as large as 0.9, the estimates of the variance of the errors (the root mean squared errors) are biased toward zero (Kiviet and Krämer 1992). When a first-order autocorrelation coefficient for the errors is estimated with the other coefficients, the procedure essentially constructs the errors and their harmonics to obtain the best fit. Because the harmonic of the constructed error is likely to resemble those of the other variables, the estimates of the coefficients for the explanatory variables may be biased (Yule 1926). This bias appeared to be especially great in the equations for structures. For further discussion of these equations see Kopcke (1993).

Quarterly Models of Investment in Equipment and Structures

Accelerator

$$IS = 23.45 + \sum_{i=1}^{11} b_i RGDPBUS_{t-i} - .20KS_{t-1}$$

b_1 = .0557
b_2 = .0318
b_3 = .0171
b_4 = .0094
b_5 = .0071
b_6 = .0080
b_7 = .0102
b_8 = .0119
b_9 = .0110
b_{10} = .0057
b_{11} = −.0060
Sum = .1618

$$IS = 67.99 + \sum_{i=1}^{11} b_i RGDPBUS_{t-1} + .03KS_{t-1} + .51EXPITOTL$$

b_1 = .0101
b_2 = −.0060
b_3 = −.0137
b_4 = −.0150
b_5 = −.0119
b_6 = −.0064
b_7 = −.0006
b_8 = .0036
b_9 = .0041
b_{10} = −.0012
b_{11} = −.0142
Sum = −.0512

$$IS = 4.76 + \sum_{i=1}^{11} b_i RGDPBUS_{t-i} - .01KS_{t-i} + .92IS_{t-1}$$

b_1 = .0196
b_2 = .0058
b_3 = −.0025
b_4 = −.0063
b_5 = −.0067
b_6 = −.0049
b_7 = −.0020
b_8 = .0009
b_9 = .0027
b_{10} = .0021
b_{11} = −.0019
Sum = .0070

$$IE = -162.16 + \sum_{i=1}^{3} b_i RGDPBUS_{t-i} - .11KE_{t-1}$$

b_1 = .1042
b_2 = .0628
b_3 = .0213
Sum = .1884

$$IE = -217.11 + \sum_{i=1}^{3} b_i RGDPBUS_{t-i}$$

$$- .18KE_{t-1} - .34EXPITOTL$$

b_1 = .1036
b_2 = .0933
b_3 = .0831
Sum = .2799

$$IE = -41.92 + \sum_{i=1}^{3} b_i RGDPBUS_{t-i} - .03KE_{t-1} + .77IE_{t-1}$$

b_1 = .0708
b_2 = .0161
b_3 = −.0387
Sum = .0482

Neoclassical (period of fit for equipment: 1962:III–1979:IV, period of fit for structures: 1964:II–1977:IV)

$$IS = 39.84 + \sum_{i=1}^{20} b_i (RGDPBUS/RS)_{t-i}$$

$$+ \sum_{i=1}^{20} c_i (RGDPBUS_{t-i}/RS_{t-1-i}) + .02KS_{t-1}$$

b_1 = .0004
b_2 = .0001
b_3 = −.0006
b_4 = −.0018
b_5 = −.0031
b_6 = −.0047
b_7 = −.0064
b_8 = −.0081
b_9 = −.0098
b_{10} = −.0114
b_{11} = −.0128
b_{12} = −.0139
b_{13} = −.0146
b_{14} = −.0149
b_{15} = −.0147
b_{16} = −.0139
b_{17} = −.0125
b_{18} = −.0103
b_{19} = −.0072
b_{20} = −.0033
Sum = −.1635

$c_1 = .0000$
$c_2 = .0009$
$c_3 = .0021$
$c_4 = .0034$
$c_5 = .0048$
$c_6 = .0064$
$c_7 = .0079$
$c_8 = .0094$
$c_9 = .0108$
$c_{10} = .0120$
$c_{11} = .0129$
$c_{12} = .0136$
$c_{13} = .0140$
$c_{14} = .0140$
$c_{15} = .0135$
$c_{16} = .0125$
$c_{17} = .0109$
$c_{18} = .0086$
$c_{19} = .0057$
$c_{20} = .0021$
Sum = .1653

$$IS = 43.86 + \sum_{i=1}^{20} b_i (RGDPBUS/RS)_{t-i}$$

$$+ \sum_{i=1}^{20} c_i (RGDPBUS_{t-i}/RS_{t-1-i}) + .01KS_{t-1} + .07EXPITOTL$$

$b_1 = .0008$
$b_2 = .0018$
$b_3 = .0020$
$b_4 = .0018$
$b_5 = .0010$
$b_6 = -.0001$
$b_7 = -.0016$
$b_8 = -.0032$
$b_9 = -.0049$
$b_{10} = -.0066$
$b_{11} = -.0082$
$b_{12} = -.0097$
$b_{13} = -.0108$
$b_{14} = -.0116$
$b_{15} = -.0119$
$b_{16} = -.0116$
$b_{17} = -.0107$
$b_{18} = -.0090$
$b_{19} = -.0064$
$b_{20} = -.0029$
Sum = −.1018

$c_1 = -.0016$
$c_2 = -.0017$
$c_3 = -.0013$
$c_4 = -.0006$
$c_5 = .0004$
$c_6 = .0017$
$c_7 = .0031$
$c_8 = .0046$
$c_9 = .0061$
$c_{10} = .0075$
$c_{11} = .0088$
$c_{12} = .0098$
$c_{13} = .0106$
$c_{14} = .0111$
$c_{15} = .0110$
$c_{16} = .0105$
$c_{17} = .0094$
$c_{18} = .0076$
$c_{19} = .0051$
$c_{20} = .0018$
Sum = .1038

$$IS = -2.21 + \sum_{i=1}^{20} b_i (RGDPBUS/RS)_{t-i}$$

$$+ \sum_{i=1}^{20} c_i (RGDPBUS_{t-i}/RS_{t-1-i}) - .00KS_{t-1} + .82IS_{t-1}$$

$b_1 = .0016$
$b_2 = .0041$
$b_3 = .0060$
$b_4 = .0073$
$b_5 = .0081$
$b_6 = .0085$
$b_7 = .0085$
$b_8 = .0082$
$b_9 = .0076$
$b_{10} = .0069$
$b_{11} = .0060$
$b_{12} = .0050$
$b_{13} = .0041$
$b_{14} = .0031$
$b_{15} = .0024$
$b_{16} = .0017$
$b_{17} = .0014$
$b_{18} = .0013$
$b_{19} = .0015$
$b_{20} = .0022$
Sum = .0954

$c_1 = -.0037$
$c_2 = -.0055$
$c_3 = -.0068$
$c_4 = -.0076$
$c_5 = -.0081$
$c_6 = -.0082$
$c_7 = -.0080$
$c_8 = -.0076$
$c_9 = -.0070$
$c_{10} = -.0062$
$c_{11} = -.0054$
$c_{12} = -.0045$
$c_{13} = -.0036$
$c_{14} = -.0027$
$c_{15} = -.0020$
$c_{16} = -.0014$
$c_{17} = -.0010$
$c_{18} = -.0009$

$c_{19} = -.0011$
$c_{20} = -.0016$
Sum = $-.0928$

$$IE = -30.37 + \sum_{i=1}^{13} b_i (RGDPBUS/RE)_{t-i} + \sum_{i=1}^{13} c_i (RGDPBUS_{t-i}/RE_{t-1-i}) + .10KE_{t-1}$$

$b_1 = .0027$
$b_2 = -.0134$
$b_3 = -.0259$
$b_4 = -.0350$
$b_5 = -.0411$
$b_6 = -.0444$
$b_7 = -.0451$
$b_8 = -.0435$
$b_9 = -.0399$
$b_{10} = -.0345$
$b_{11} = -.0275$
$b_{12} = -.0193$
$b_{13} = -.0100$
Sum = $-.3767$

$c_1 = .0157$
$c_2 = .0264$
$c_3 = .0345$
$c_4 = .0401$
$c_5 = .0433$
$c_6 = .0443$
$c_7 = .0432$
$c_8 = .0402$
$c_9 = .0353$
$c_{10} = .0288$
$c_{11} = .0208$
$c_{12} = .0113$
$c_{13} = .0006$
Sum = $.3844$

$$IE = -34.00 + \sum_{i=1}^{13} b_i (RGDPBUS/RE)_{t-i} + \sum_{i=1}^{13} c_i (RGDPBUS_{t-i}/RE_{t-1-i}) + .10KE_{t-1} + .18EXPITOTL$$

$b_1 = .0026$
$b_2 = -.0124$
$b_3 = -.0236$
$b_4 = -.0314$
$b_5 = -.0363$
$b_6 = -.0384$
$b_7 = -.0382$
$b_8 = -.0360$
$b_9 = -.0320$
$b_{10} = -.0267$
$b_{11} = -.0204$
$b_{12} = -.0133$
$b_{13} = -.0059$
Sum = $-.3120$

$c_1 = .0139$
$c_2 = .0236$
$c_3 = .0306$
$c_4 = .0351$
$c_5 = .0373$
$c_6 = .0374$
$c_7 = .0357$
$c_8 = .0323$
$c_9 = .0276$
$c_{10} = .0216$
$c_{11} = .0147$
$c_{12} = .0069$
$c_{13} = -.0013$
Sum = $.3154$

$$IE = -10.90 + \sum_{i=1}^{13} b_i (RGDPBUS/RE)_{t-i} + \sum_{i=1}^{13} c_i (RGDPBUS_{t-i}/RE_{t-1-i}) + .04KE_{t-1} + .66IE_{t-1}$$

$b_1 = .0036$
$b_2 = -.0057$
$b_3 = -.0123$
$b_4 = -.0163$
$b_5 = -.0182$
$b_6 = -.0184$
$b_7 = -.0171$
$b_8 = -.0148$
$b_9 = -.0118$
$b_{10} = -.0084$
$b_{11} = -.0050$
$b_{12} = -.0020$
$b_{13} = -.0003$
Sum = $-.1261$

$c_1 = .0060$
$c_2 = .0118$
$c_3 = .0157$
$c_4 = .0176$
$c_5 = .0178$
$c_6 = .0168$
$c_7 = .0147$
$c_8 = .0120$
$c_9 = .0088$
$c_{10} = .0057$
$c_{11} = .0028$
$c_{12} = .0005$
$c_{13} = -.0009$
Sum = $.1288$

q Model

$$IS = 1.70 + \sum_{i=1}^{8} b_i (q-1)_{t-1} KS_{t-1-i} + .09KS_{t-1}$$

$b_1 = -.0022$
$b_2 = .0080$
$b_3 = .0120$
$b_4 = .0116$
$b_5 = .0086$
$b_6 = .0050$
$b_7 = .0025$
$b_8 = .0030$
Sum = .0484

$$IS = 32.74 + \sum_{i=1}^{8} b_i (q-1)_{t-i} KS_{t-1-i} + .02KS_{t-1} + .28EXPITOTL$$

$b_1 = .0025$
$b_2 = .0063$
$b_3 = .0065$
$b_4 = .0046$
$b_5 = .0018$
$b_6 = -.0004$
$b_7 = -.0006$
$b_8 = .0026$
Sum = .0232

$$IS = 1.04 + \sum_{i=1}^{8} b_i (q-1)_{t-i} KS_{t-1-i} + .01KS_{t-1} + .88IS_{t-1}$$

$b_1 = .0026$
$b_2 = .0049$
$b_3 = .0044$
$b_4 = .0021$
$b_5 = -.0008$
$b_6 = -.0030$
$b_7 = -.0033$
$b_8 = -.0007$
Sum = .0062

$$IE = -30.57 + \sum_{i=1}^{5} b_i (q-1)_{t-i} KE_{t-1-i} + .21KE_{t-1}$$

$b_1 = -.0346$
$b_2 = .0324$
$b_3 = .0321$
$b_4 = .0085$
$b_5 = .0057$
Sum = .0441

$$IE = -8.87 + \sum_{i=1}^{5} b_i (q-1)_{t-i} KE_{t-1-i} - .02KE_{t-1} + .78EXPITOTL$$

$b_1 = -.0276$
$b_2 = .0141$
$b_3 = .0085$
$b_4 = -.0178$
$b_5 = -.0383$
Sum = -.0610

$$IE = -3.19 + \sum_{i=1}^{5} b_i (q-1)_{t-i} KE_{t-1-i} + .01KE_{t-1} + 1.01IE_{t-1}$$

$b_1 = .0159$
$b_2 = .0177$
$b_3 = .0024$
$b_4 = -.0146$
$b_5 = -.0180$
Sum = .0035

Cash Flow

$$IS = 25.32 + \sum_{i=1}^{12} b_i (F/CS)_{t-i}$$

$b_1 = .1297$
$b_2 = .0721$
$b_3 = .0338$
$b_4 = .0117$
$b_5 = .0024$
$b_6 = .0027$
$b_7 = .0094$
$b_8 = .0191$
$b_9 = .0286$
$b_{10} = .0347$
$b_{11} = .0342$
$b_{12} = .0237$
Sum = .4020

$$IS = 55.20 + \sum_{i=1}^{12} b_i (F/CS)_{t-i} + .32EXPITOTL$$

$b_1 = -.0029$
$b_2 = -.0023$
$b_3 = -.0025$
$b_4 = -.0034$
$b_5 = -.0047$
$b_6 = -.0062$
$b_7 = -.0076$
$b_8 = -.0086$
$b_9 = -.0091$
$b_{10} = -.0087$
$b_{11} = -.0072$
$b_{12} = -.0044$
Sum = -.0675

$$IS = 4.26 + \sum_{i=1}^{12} b_i (F/CS)_{t-i} + .93IS_{t-1}$$

$b_1 = .0475$
$b_2 = .0173$
$b_3 = -.0026$

$b_4 = -.0140$
$b_5 = -.0184$
$b_6 = -.0176$
$b_7 = -.0130$
$b_8 = -.0064$
$b_9 = .0006$
$b_{10} = .0064$
$b_{11} = .0093$
$b_{12} = .0077$
Sum = .0167

$$IE = -25.34 + \sum_{i=1}^{5} b_i (F/CE)_{t-i}$$

$b_1 = .6998$
$b_2 = .0661$
$b_3 = .0440$
$b_4 = .1776$
$b_5 = .0112$
Sum = .9987

$$IE = -42.84 + \sum_{i=1}^{5} b_i (F/CE)_{t-i} + .43EXPITOTL$$

$b_1 = .5892$
$b_2 = -.0356$
$b_3 = -.0016$
$b_4 = .1456$
$b_5 = -.1395$
Sum = .5580

$$IE = 1.03 + \sum_{i=1}^{5} b_i (F/CE)_{t-i} + .98IE_{t-1}$$

$b_1 = 0.1939$
$b_2 = -.0921$
$b_3 = -.0590$
$b_4 = .0370$
$b_5 = -.0600$
Sum = .0198

Autoregression

$$IS = 6.00 + \sum_{i=1}^{4} b_i IS_{t-i}$$

$b_1 = 1.1277$
$b_2 = .0587$
$b_3 = -.1832$
$b_4 = -.0516$
Sum = .9516

$$IS = 12.64 \sum_{i=1}^{4} b_i IS_{t-i} + .06EXPITOTL$$

$b_1 = 1.008$
$b_2 = .0518$
$b_3 = -.1405$
$b_4 = -.1425$
Sum = .7768

$$IS = 4.62 + \sum_{i=1}^{4} b_i IS_{t-i} + \sum_{i=1}^{3} c_i RGDPBUS_{t-i}$$

$b_1 = .8418$
$b_2 = .1980$
$b_3 = .0624$
$b_4 = -.1619$
Sum = .9403

$c_1 = .0323$
$c_2 = .0002$
$c_3 = -.0319$
Sum = .0006

$$IS = 16.21 + \sum_{i=1}^{4} b_i IS_{t-i} + \sum_{i=1}^{3} c_i RGDPBUS_{t-i} + .10EXPITOTL$$

$b_1 = .7871$
$b_2 = .1156$
$b_3 = .0398$
$b_4 = -.1739$
Sum = .7686

$c_1 = .0241$
$c_2 = .0005$
$c_3 = -.0306$
Sum = −.0060

$$IE = 2.68 + \sum_{i=1}^{4} b_i IE_{t-i}$$

$b_1 = 1.2674$
$b_2 = .0369$
$b_3 = -.3013$
$b_4 = -.0128$
Sum = .9902

$$IE = 1.28 + \sum_{i=1}^{4} b_i IE_{t-i} + .02EXPITOTL$$

$b_1 = 1.2519$
$b_2 = .0367$
$b_3 = -.3022$
$b_4 = -.0195$
Sum = .9669

$$IE = -28.50 + \sum_{i=1}^{4} b_i IS_{t-i} + \sum_{i=1}^{2} c_i RGDPBUS_{t-i}$$

$b_1 = .9052$
$b_2 = .1422$

$b_3 = -.1363$
$b_4 = -.1966$
Sum $= .7145$

$c_1 = .0369$
$c_2 = -.0040$
Sum $= .0329$

$$IE = -28.12 + \sum_{i=1}^{4} b_i IS_{t-i} + \sum_{i=1}^{2} c_i RGDPBUS_{t-i} - .03EXPITOTL$$

$b_1 = .9149$
$b_2 = .1407$
$b_3 = -.1292$
$b_4 = -.1958$
Sum $= .7306$

$c_1 = .0365$
$c_2 = -.0022$
Sum $= .0343$

Annual Models of Total Investment

Accelerator

$$IS = 30.68 + .11RGDPBUS_{t-1} - .13KS_{t-1}$$

$$IS = 71.30 - .04RGDPBUS_{t-1} + .01KS_{t-1} + .50YRAHEAD$$

$$IS = 24.16 + .06RGDPBUS_{t-1} - .08KS_{t-1} + 0.46IS_{t-1}$$

$$IE = -133.18 + .16RGDPBUS_{t-1} - .07KE_{t-1}$$

$$IE = -157.68 + .20RGDPBUS_{t-1} - .10KE_{t-1} - .15YRAHEAD$$

$$IE = -123.22 + .16RGDPBUS_{t-1} - .03KE_{t-1} - .34IE_{t-1}$$

Neoclassical (Period of fit for Structures: 1964–1977. Period of fit for Equipment: 1964–1979)

$$IS = 67.01 + \sum_{i=1}^{4} b_i (RGDPBUS/RS)_{t-i} + \sum_{i=1}^{4} c_i (RGDPBUS_{t-i}/RS_{t-1-i}) + .02KS_{t-1}$$

$b_1 = -.0004$
$b_2 = -.0100$
$b_3 = -.0235$
$b_4 = -.0307$
Sum $= -.0645$

$c_1 = .0091$
$c_2 = .0160$
$c_3 = .0286$
$c_4 = .0100$
Sum $= .0638$

$$IS = 96.30 + \sum_{i=1}^{4} b_i (RGDPBUS/RS)_{t-i} + \sum_{i=1}^{4} c_i (RGDPBUS_{t-i}/RS_{t-1-i}) - .01KS_{t-1} + .22YRAHEAD$$

$b_1 = -.0009$
$b_2 = -.0032$
$b_3 = -.0202$
$b_4 = -.0290$
Sum $= -.0534$

$c_1 = .0024$
$c_2 = .0124$
$c_3 = .0265$
$c_4 = .0090$
Sum $= .0503$

$$IS = 74.11 + \sum_{i=1}^{4} b_i (RGDPBUS/RS)_{t-i} + \sum_{i=1}^{4} c_i (RGDPBUS_{t-i}/RS_{t-1-i}) + .04KS_{t-1} - .22IS_{t-1}$$

$b_1 = -.0014$
$b_2 = -.0165$
$b_3 = -.0336$
$b_4 = -.0370$
Sum $= -.0886$

$c_1 = .0146$
$c_2 = .0253$
$c_3 = .0356$
$c_4 = .0118$
Sum $= .0873$

$$IE = -29.51 + \sum_{i=1}^{4} b_i (RGDPBUS/RE)_{t-i} + \sum_{i=1}^{4} c_i (RGDPBUS_{t-i}/RE_{t-1-i}) + .08KE_{t-1}$$

$b_1 = .0135$
$b_2 = -.0467$
$b_3 = -.0335$
$b_4 = -.0177$
Sum $= -.0843$

$c_1 = .0322$
$c_2 = .0389$
$c_3 = .0173$
$c_4 = .0079$
Sum = .0963

$$IE = -30.33 + \sum_{i=1}^{4} b_i(RGDPBUS/RE)_{t-i}$$

$$+ \sum_{i=1}^{4} c_i(RGDPBUS_{t-i}/RE_{t-1-i}) + .08KE_{t-1} + .02YRAHEAD$$

$b_1 = .0135$
$b_2 = -.0465$
$b_3 = -.0330$
$b_4 = -.0162$
Sum = −.0822

$c_1 = .0317$
$c_2 = .0388$
$c_3 = .0158$
$c_4 = .0075$
Sum = .0938

$$IE = -45.52 + \sum_{i=1}^{4} b_i(RGDPBUS/RE)_{t-i}$$

$$+ \sum_{i=1}^{4} c_i(RGDPBUS_{t-i}/RE_{t-1-i}) + .14KE_{t-1} - .68IE_{t-1}$$

$b_1 = .0113$
$b_2 = -.0536$
$b_3 = -.0517$
$b_4 = -.0451$
Sum = −.1392

$c_1 = .0481$
$c_2 = .0522$
$c_3 = .0441$
$c_4 = .0120$
Sum = .1564

q Model

$$IS = 6.99 + .02((q-1) * KS_{t-1})_{t-1} - .08KS_{t-1}$$

$$IS = 74.93 - .02((q-1) * KS_{t-1})_{t-1} - .07KS_{t-1} + .54YRAHEAD$$

$$IS = 8.43 + .00((q-1) * KS_{t-1})_{t-1} + .02KS_{t-1} + .70IS_{t-1}$$

$$IE = -33.06 + .05((q-1) * KE_{t-1})_{t-1} + .22KE_{t-1}$$

$$IE = -2.69 - .07((q-1) * KE_{t-1})_{t-1} - .02KE_{t-1} + .76YRAHEAD$$

$$IE = -16.71 + .03((q-1) * KE_{t-1})_{t-1} + .11KE_{t-1} + .52IE_{t-1}$$

Cash Flow

$$IS = 33.99 + \sum_{i=1}^{2} b_i(F/CS)_{t-i}$$

$b_1 = .2607$
$b_2 = .1050$
Sum = .3656

$$IS = 58.84 + \sum_{i=1}^{2} b_i(F/CS)_{t-i} + .33YRAHEAD$$

$b_1 = -.0291$
$b_2 = -.0657$
Sum = −.0947

$$IS = 20.87 + \sum_{i=1}^{2} b_i(F/CS)_{t-i} + .74IS_{t-1}$$

$b_1 = .1312$
$b_2 = -.0820$
Sum = .0492

$$IE = -14.23 + \sum_{i=1}^{2} b_i(F/CE)_{t-i}$$

$b_1 = 1.0725$
$b_2 = -.1036$
Sum = .9689

$$IE = -37.04 + \sum_{i=1}^{2} b_i(F/CE)_{t-i} + .52YRAHEAD$$

$b_1 = .6595$
$b_2 = -.2237$
Sum = .4358

$$IE = 3.50 + \sum_{i=1}^{2} b_i(F/CE)_{t-i} + .81IE_{t-1}$$

$b_1 = .4601$
$b_2 = -.2745$
Sum = .1856

Autoregression

$$IS = 25.95 + \sum_{i=1}^{3} b_i IS_{t-i}$$

$b_1 = 1.1330$
$b_2 = -.8712$

b_3 = .5391
Sum = .8009

$$IS = 52.44 + \sum_{i=1}^{3} b_i IS_{t-i} + .31YRAHEAD$$

b_1 = .2204
b_2 = −.2408
b_3 = −.0667
Sum = −.0871

$$IS = 24.92 + \sum_{i=1}^{3} b_i IS_{t-i} + \sum_{i=1}^{2} c_i RGDPBUS_{t-i}$$

b_1 = .9920
b_2 = −.5068
b_3 = .6344
Sum = 1.1196

c_1 = .0279
c_2 = −.0441
Sum = −.0162

$$IS = 65.70 + \sum_{i=1}^{3} b_i IS_{t-i} + \sum_{i=1}^{2} c_i RGDPBUS_{t-i} + .43YRAHEAD$$

b_1 = .1851
b_2 = −.2139
b_3 = .2098
Sum = .1810

c_1 = −.0266
c_2 = −.0047
Sum = −.0313

$$IE = 7.69 + \sum_{i=1}^{3} b_i IE_{t-i}$$

b_1 = 1.4548
b_2 = −1.2027
b_3 = .7796
Sum = 1.0317

$$IE = -1.85 + \sum_{i=1}^{3} b_i IE_{t-i} + .15YRAHEAD$$

b_1 = 1.2625
b_2 = −1.0930
b_3 = .6951
Sum = .8646

$$IE = -173.04 \sum_{i=1}^{3} b_i IE_{t-i} + \sum_{i=1}^{2} c_i RGDPBUS_{t-i}$$

b_1 = −.9458
b_2 = .7176
b_3 = −.5005
Sum = −0.7287

c_1 = .3993
c_2 = −.2056
Sum = .1937

$$IE = -174.94 + \sum_{i=1}^{3} b_i IE_{t-i} + \sum_{i=1}^{2} c_i RGDPBUS_{t-i} - .04YRAHEAD$$

b_1 = −.9630
b_2 = .7540
b_3 = −.5143
Sum = −.7233

c_1 = .4128
c_2 = −.2143
Sum = .1985

Richard W. Kopcke, "The Determinants of Business Investment: Has Capital Spending Been Surprisingly Low?" *New England Economic Review*, January/February 1993, pp. 3–31.

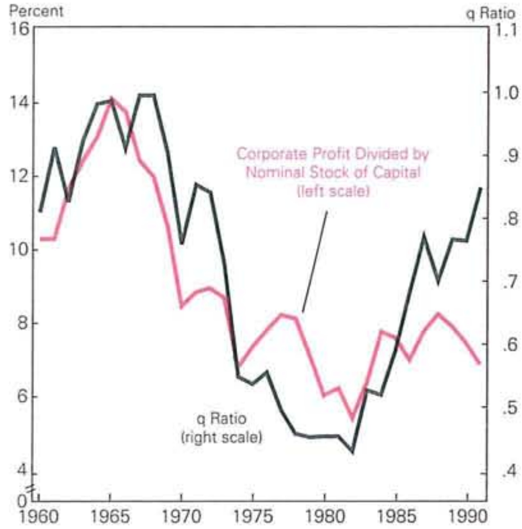
ERRATA

Three figures appeared with incorrect color keys in this article. The corrected figures are presented on both sides of this page.

Figure 2 (page 13)

Figure 2

Rate of Return on the Stock of Capital and the q Ratio (q Model)

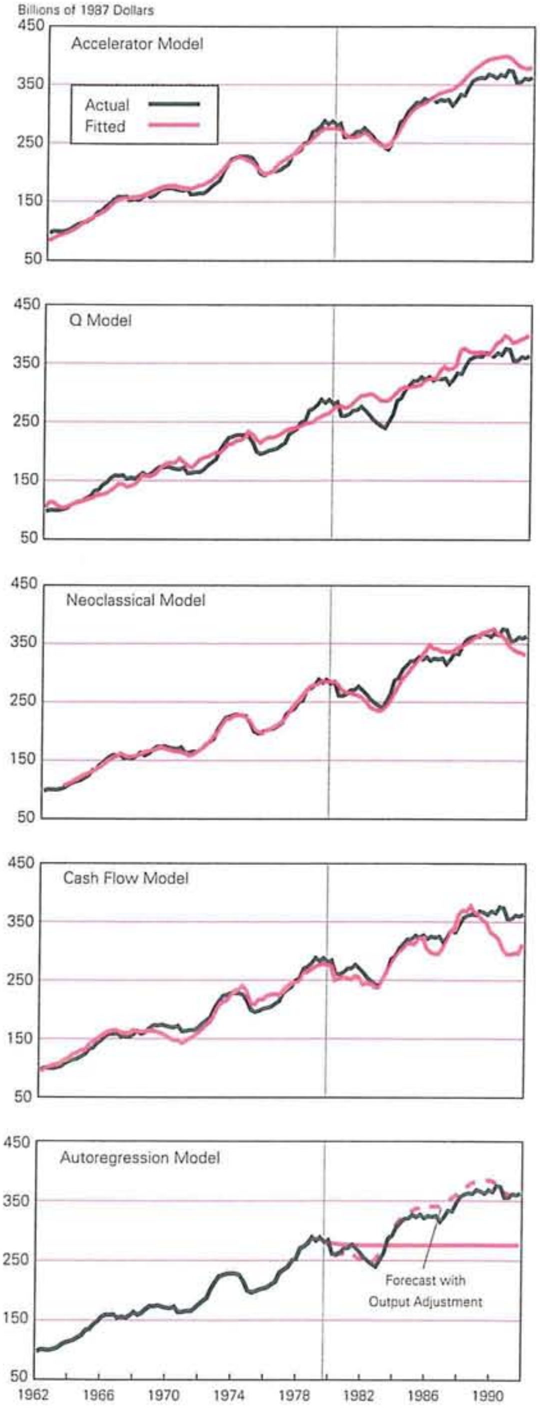


Source of data: Data Resources, Inc.

See over.

Figure 5 (page 21)

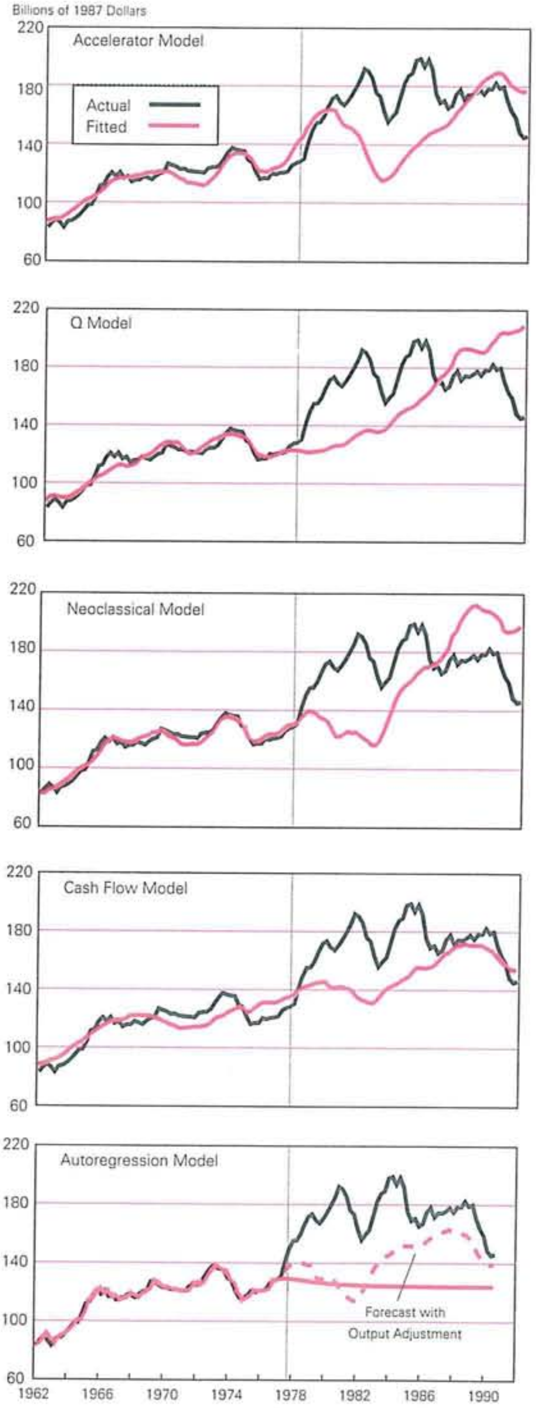
Actual Expenditures and Forecasts[a] of Investment in Equipment



[a]Model forecasts for the period 1980:I to 1992:I; model estimates fit to actual data through 1979:IV.
Source of data: Data Resources, Inc.

Figure 6 (page 23)

Actual Expenditures and Forecasts[a] of Investment in Structures



[a]Model forecasts for the period 1978:I to 1992:I; model estimates fit to actual data through 1977:IV.
Source of data: Data Resources, Inc.

Federal Reserve Bank of Boston
P.O. Box 2076
Boston, Massachusetts 02106-2076

Address Correction Requested

Bulk Rate
U.S. Postage
PAID
Richmond, VA
Permit No. 930