FILE COPY

# New England Economic Review

Federal Reserve Bank of Boston

Sept./Oct. 1996

*Katerina Simons* — *Value at Risk—New Approaches to Risk Management*

*Lynn Elaine Browne with Joshua Gleason* — *The Saving Mystery, or Where Did the Money Go?*

*Peter Fortune* — *Do Municipal Bond Yields Forecast Tax Policy?*

*Joe Peek and Eric S. Rosengren* — *The Use of Capital Ratios to Trigger Intervention in Problem Banks: Too Little, Too Late*

**President**
Cathy E. Minehan

**Senior Vice President and Director of Research**
Lynn Elaine Browne

*Research Department:*
**Banking and Monetary Policy**
Eric S. Rosengren, Vice President
John S. Jordan
Joe Peek
Katerina Simons
Robert Tannenwald

**International**
Norman S. Fieleke, Vice President

**Macro/Finance**
Richard W. Kopcke, Vice President
Peter Fortune
Geoffrey M. B. Tootell, Assistant Vice President

**Open Economy Macroeconomics**
Jeffrey C. Fuhrer, Vice President
Jane Sneddon Little, Assistant Vice President

**Regional**
Katharine L. Bradbury, Vice President
Yolanda K. Kodrzycki
Christopher J. Mayer
Joanna Stavins
Robert K. Triest

The **New England Economic Review** (ISSN 0028-4726) is a publication of the Research Department of the Federal Reserve Bank of Boston. Joan T. Poskanzer is the editor. Ann Eggleston coordinates the design and production. 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

**Value at Risk—New Approaches to Risk Management**

*Katerina Simons*

Managing risk has always been an integral part of banking. In the past two years an approach to risk management called "Value at Risk" has been accepted by both practitioners and regulators as the "right" way to measure risk, becoming a de facto industry standard. Yet, the danger is that overreliance on value at risk can give risk managers a false sense of security or lull them into complacency. Value at risk is only one of many tools of managing risk, and it is based on a number of unrealistic assumptions. There is no generally accepted way to calculate it, and various methods can yield widely different results.

This article describes several common methods for calculating value at risk and highlights important assumptions and methodological issues. The author discusses the strengths and weaknesses of value at risk, pointing out that its use has created a common language for discussions about risk and prompted more dialogue about risk issues. She cautions, however, that successful risk management is a much broader task, which depends crucially on appropriate incentives and internal controls. 3

**The Saving Mystery, or Where Did the Money Go?**

*Lynn Elaine Browne with Joshua Gleason*

Of great concern and puzzlement to many has been the decline in the U.S. personal saving rate. From 8 percent of personal income 20 years ago, saving has fallen to less than 4 percent. This is a matter of concern because saving and investment are closely linked, and investment is believed critical to productivity gains and a rising standard of living. The decline in saving is also a source of puzzlement because it runs counter to many people's perception of what is happening.

This article investigates the decline in saving, focusing on "where the money went." The authors find that rising expenditures on medical services are absorbing a growing fraction of income. Thus, the saving problem is not about thrift versus profligacy, but rather a competition between more and better medical care, on the one hand, and more investment, on the other. They point out that efforts to stimulate saving are only one way to increase the economy's productive capacity, and that the ultimate goal is higher standards of living. 15

# Contents

**Do Municipal Bond Yields Forecast Tax Policy?**

*Peter Fortune*

During the recent flat tax debate, interest rates on long-term municipal bonds rose relative to the rate on U.S. Treasury bonds. This was widely attributed to expectations of a reduction in future tax rates. While an axiom of finance states that current asset prices reflect expectations about future events, there is no consensus on how sensitive municipal bond yields are to expectations about future tax rates. This study assesses that question by examining the relationship between the implicit tax rate and actual future tax rates.

Efficient markets theory predicts that the implicit tax rate—the tax rate that equates the after-tax yield on a Treasury bond to the yield on a tax-exempt bond—will be an excellent predictor of future tax rates. The author finds that although the Efficient Markets prediction is not supported, implicit tax rates do contain some information about future tax rates. The information content in implicit tax rates is particularly high around the time of major tax debates that have resulted in significant changes in tax rates. At other times, including the flat tax debate of 1995–96, implicit tax rates carry little information about future tax rates. 29

**The Use of Capital Ratios to Trigger Intervention in Problem Banks: Too Little, Too Late**

*Joe Peek and Eric S. Rosengren*

A wave of depository institution failures and dramatic losses to deposit insurance funds occurred in the 1980s and continued into the 1990s. In response, the Congress passed a series of bank regulatory acts intended to address the problems that led to the crisis and prevent its recurrence. The Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA) was the capstone of this transformation of banking legislation, with two key provisions designed to reduce the cost of troubled banks to the deposit insurance fund: early closure of failing institutions, and early supervisory intervention in problem banks, referred to as prompt corrective action.

This article considers whether the capital ratio thresholds that trigger prompt corrective action intervention provide sufficient lead time for successful intervention at troubled banks. The study finds that because prompt corrective action is based on a lagging indicator of a bank's financial health, it is likely to trigger intervention in problem banks only after they have been identified by examiners, who rely on far more information than the capital ratio. The authors propose two simple measures to improve the current triggers for prompt corrective action. 49

# Value at Risk—New Approaches to Risk Management

Katerina Simons

*Economist, Federal Reserve Bank of Boston.* *The author thanks Peter Fortune, Richard Kopcke, and James O'Brien for helpful comments and Timothy Lin for able research assistance.*

Managing risk has always been an integral part of banking. Recently, however, "risk management" has become a popular buzzword—the phrase appeared in the *American Banker* 72 times in 1990 and 325 times in 1995.[1] At the center of the recent interest is an approach to risk management called "Value at Risk." In the past two years it has been accepted by both practitioners and regulators as the "right" way to measure risk, becoming a de facto industry standard. Yet, the danger is that overreliance on value at risk can give risk managers a false sense of security or lull them into complacency. After all, value at risk is only one of many tools for managing risk, and it is based on a number of unrealistic assumptions. Moreover, there is no generally accepted way to calculate value at risk, and various methods can yield widely different results.

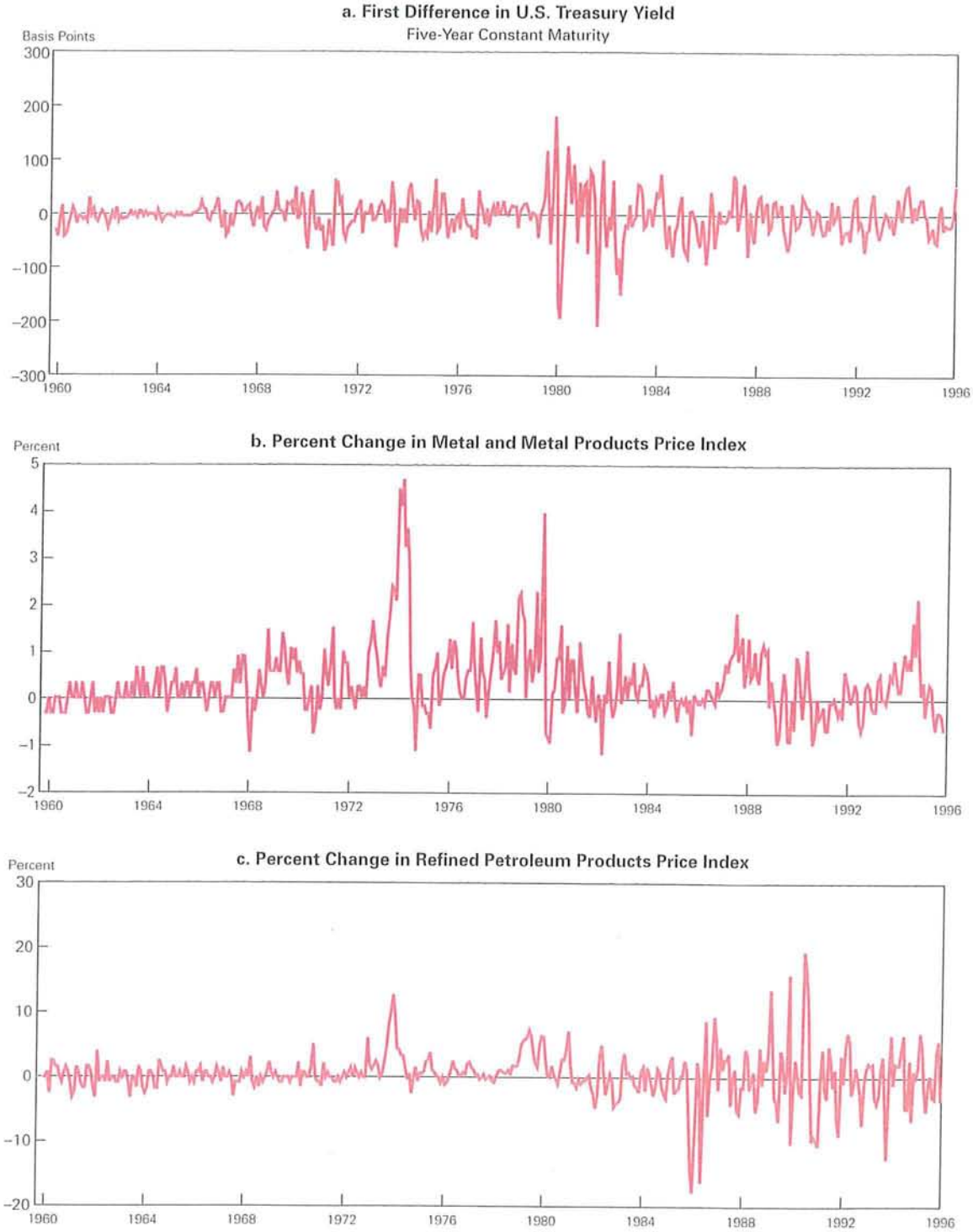
This article will review briefly the reasons for the new approaches and describe the Basle Market Risk Standard, which proposes the use of banks' internal value-at-risk models to set appropriate capital levels to cover market risk in bank trading operations. The article will describe several common methods for calculating value at risk (VAR) and highlight important assumptions and methodological issues. These issues will be illustrated by two step-by-step examples of calculating VAR for a single instrument. The article concludes with a brief discussion of strengths and weaknesses of VAR.

## I. Why a New Approach?

Increased volatility in the financial markets since the 1970s has spurred new emphasis on risk management. Increased volatility first became apparent in the currency markets after the collapse of the Bretton Woods Agreement, followed, in short order, by interest rates and

Figure 1

*Examples of Volatility Patterns*



commodity prices. Figures 1a, 1b, and 1c depict examples of these volatility patterns.

Rapid advances in information technology have increased proprietary trading activity and heightened the emphasis on money management performance. At the same time, the growing complexity of financial products, particularly derivatives, has made it more difficult to evaluate and measure the risks taken by financial institutions, as accounting and disclosure rules have failed to keep pace with financial innovation. The use of derivatives has also increased linkages between markets. For example, a shock in the equity futures market of one country can be transmitted rapidly to the market for the underlying equities and perhaps to currency and equity markets of other countries as well.

*The growing complexity of financial products, particularly derivatives, has made it more difficult to evaluate and measure the risks taken by financial institutions, as accounting and disclosure rules have failed to keep pace with financial innovation.*

The sharp rise in transaction volume in derivatives markets, coupled with several well-publicized trading losses, has led to a new interest in an "objective" way of gauging the adequacy of capital. In their search, financial institutions turned in part to analytical tools introduced in derivatives markets, and VAR emerged as the favored method for measuring risk. The 1993 study by the Group of Thirty, *Derivatives: Practices and Principles*, strongly recommends VAR analysis; that study's recommendations have been broadly accepted by the industry as the standard of "best practices."

Currently, derivatives techniques have spread to many instruments and their structures have become increasingly complex. More than 1,000 banks, nonfinancial corporations, insurance companies, mutual funds, and other asset managers use them to manage their risks. The availability of data on derivatives prices from the past decade gives a better empirical foundation to VAR analysis. The Bank for International Settlements (BIS) has allowed banks to use their own internal models of risk in setting capital requirements for market risk. The acceptable models must rely on VAR methodologies.

VAR has an intuitive appeal because it summarizes the risk of the entire portfolio in a single number. Moreover, it expresses in dollar terms the major concern of risk management—the potential loss to portfolio value. VAR can be applied to many different instruments and can calculate and aggregate risk across instruments and types of assets.

VAR is applied primarily to market risk, though applications have recently been expanded to incorporate credit risk. (See the box for definitions of the main types of risks in banking.) VAR holds promise of combining all quantifiable risks across the business lines of an institution, yielding one firm-wide measure of risk.

## II. What Is Value at Risk?

Essentially, VAR poses the question: "How much money might we lose over the next period of time?" Rephrasing it more precisely, "Over a given period of time with a given probability, how much could the value of the portfolio decline?" For example, if the given time is one week, the given probability 1 percent, and the value at risk $20 million, then we estimate that the odds that this portfolio will decline in value by more than $20 million within the next week are 1 in a 100.

To calculate VAR, one needs to choose a common measurement unit, a time horizon, and a probability. The common unit can be U.S. dollars, German marks, or whatever currency the organization primarily uses to do business. The chosen probability of loss usually ranges between 1 and 5 percent. The time horizon can be of any length, but it is assumed that the portfolio composition does not change during the holding period. The most common holding periods used are one day, one week, or two weeks. The choice of the holding period depends on the liquidity of the assets in the portfolio and how frequently they are traded. Relatively less liquid assets call for a longer holding period.

[1] See "New Risk Tests Win Fans—But Will They Work?" *American Banker*, 2/7/96.

**Banking—What Are the Risks?**

The risks of banking can be divided into five categories: credit, liquidity, operational, legal, and market (Figure 2).

*Credit risk* is the possibility of loss as a result of default, such as when a customer defaults on a loan, or more generally, any type of financial contract.

*Liquidity risk* is the possibility that a firm will be unable to fund its illiquid assets.

*Operational risk* is the possibility of loss resulting from errors in instructing payments or settling transactions.

*Legal risk* is the possibility of loss when a contract cannot be enforced—for example, because the customer had no authority to enter into the contract or the contract turns out to be unenforceable in a bankruptcy.

*Market risk* is the possibility of loss over a given period of time related to uncertain movements in market risk factors, such as *interest rates, currencies, equities, and commodities*. The market risk of a financial instrument can be caused by more than one factor. For example, holding a bond denominated in a foreign currency exposes one to currency risk and to interest rate risk. Similarly, entering into a domestic equity swap exposes one to equity risk and interest rate risk.

Banks may be exposed to some equity and commodity risk through swap positions and many large banks have currency risk through their currency trading, but by far the largest market risk of the banking industry is *interest rate risk*. A principal source of earnings for banks is net interest income, the difference between interest received and interest paid. The main source of interest rate risk is the volatility of those interest rates and the mismatch in the timing when the rates on assets and liabilities are reset.

Interest rate risk, in turn, can be divided into three types—*yield curve level risk, yield curve shape risk*, and *basis risk*, depending on the type of interest rate change that can cause losses.

*Yield curve level risk* refers to an equal change in rates across all maturities: for example, if all Treasury yields, from 3-month T-bills to 30-year long bonds, move up or down uniformly by 1 percent.

*Yield curve shape risk* refers to changes in the relative rates for instruments of different maturities. For example, the yield curve could "bulge," so that yields rise for intermediate maturities, such as 3- to 5-year Treasury notes, while rates for bills and long-term bonds remain unchanged.

*Basis risk* refers to the risk of changes in rates for instruments with the same maturity but pegged to a different index. For example, suppose a bank funds itself by borrowing at a 6-month Libor (London Interbank Offer Rate), a commonly used rate for interbank borrowing, and invests in an instrument whose rate is tied to the 6-month Treasury rate. If the Libor rises relative to the Treasury rate, the bank will lose money.

Of course, in reality all interest rates continuously change, exposing the bank to all three types of interest rate risk. However, it is useful to distinguish among them conceptually to pinpoint the areas of particular vulnerability.

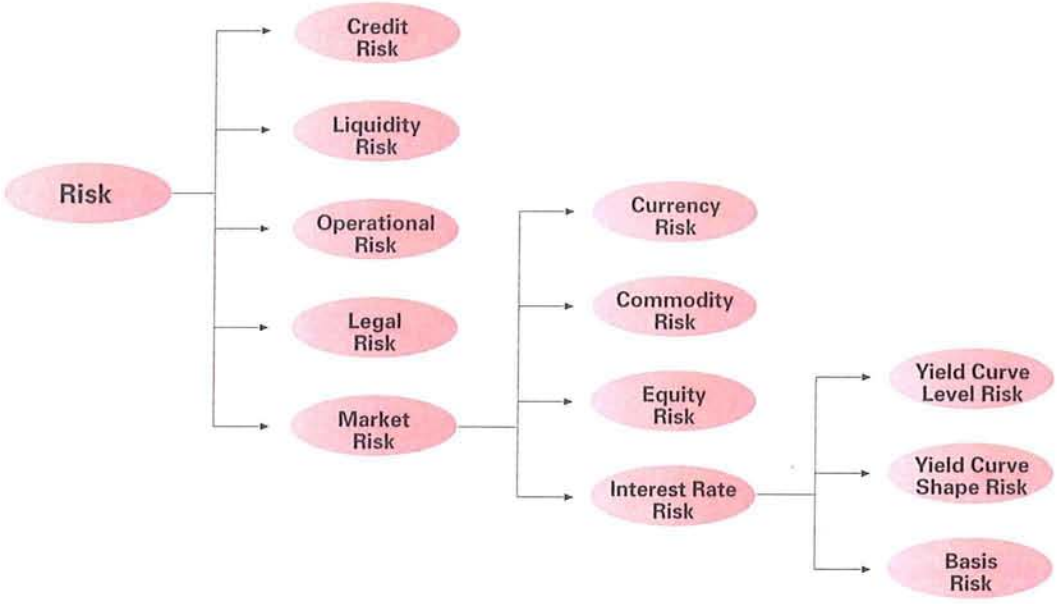
## *III. VAR and Capital Requirements for Market Risk*

VAR models have been accepted by both practitioners and bank regulators as the state of the art in quantitative risk measurement. In its recent risk-based capital proposal, the Basle Committee on Banking Supervision endorsed the use of banks' VAR models to allocate capital for market risk. The Basle standard covers internationally active banks and applies only to their trading account. The proposal offers two alternatives: "standardized" and "internal models." U.S. bank regulators favor the internal models approach, whereby the bank's own VAR model is used to set aside capital for market risk. The proposal has an implementation period of two years and will take effect in January 1998.

To be acceptable to regulators for the purposes of allocating capital, banks' internal models must meet certain qualitative and quantitative standards. In essence, qualitative standards relate to the institution's risk management function as a whole. They call for independent validation of the models by the bank or a third party; strong controls over inputs, data, and model changes; independence of the risk management function from business lines; full integration of the

Figure 2

*Types of Risk*




model into risk management; and, most important, director and senior management oversight of the risk management process.

Quantitative standards relate to specific features of the VAR model. They call for the use of a 1 percent probability level and a two-week holding period. In addition, the VAR thus found is to be multiplied by a factor of three. The multiplication factor is designed to allow for potential weaknesses in the modeling process and other nonquantifiable factors, such as incorrect assumptions about distributions, unstable volatilities, and extreme market movements.

Many practitioners, however, consider these standards too restrictive. They note that a holding period of two weeks is too long for many instruments, as traders get in and out of positions many times during a typical day. Moreover, a two-week holding period combined with a 1 percent probability safeguards against events that can be expected to occur only once in four years. This makes it difficult to validate the model within a reasonable period of time.

It should be noted that a few features of the proposal have been modified as a result of industry criticism. In particular, an earlier version of the proposal allowed the models to account for correlations of asset returns within, but not among, asset classes, such as equities, currencies, and bonds. Now, all correlations are allowed.

## *IV. Parametric VAR*

No consensus has been reached on the best way to implement VAR analysis. Most methodological issues revolve around estimation of the statistical distributions of asset returns. The main approaches are known as parametric (also known as the analytical or correlation method), historical, historical simulation, and stochastic simulation (also known as Monte Carlo).

Parametric VAR is based on the estimate of the variance–covariance matrix of asset returns, using historical time series of asset returns to calculate their standard deviations and correlations. The main assumption of the parametric VAR is that the distributions of asset returns are normal. This means that the variance–covariance matrix completely describes the distribution.[2]

[2] While a number of distributions other than normal can be completely described by their parameters, in VAR analysis the parametric approach usually refers to a normal distribution.

The parametric approach can be summarized by the equation:

$$\sigma_p^2 = \sum (a_i \cdot \sigma_i)^2 + \sum_{i \neq j} \sum_{i \neq j} a_i \cdot a_j \cdot \rho_{ij} \cdot \sigma_i \cdot \sigma_j,$$

where:

$\sigma_p^2$ = the volatility of portfolio returns,
$a_i$ = the dollar amount of the portfolio share of the $i$th instrument,
$\sigma_i^2$ = the volatility of the $i$th instrument, and
$\rho_{ij}$ = the correlation between the returns of the $i$th and $j$th instruments.

The equation shows that portfolio risk, as expressed by its variance, is a function of the variance of the return on each instrument in the portfolio, as well as on the correlations between each pair of returns. This means that unless the returns in the portfolio are perfectly correlated (all $\rho_{ij} = 1$), the variance of the portfolio does not equal the simple sum of the variances of the individual positions. When the risk that any investment contributes to the portfolio is less than the risk of that investment alone because of diversification, the risk of the portfolio is less than the sum of the risks of its parts.

The best-known parametric VAR model is J.P. Morgan's RiskMetrics. J.P. Morgan has done much to advance the public understanding and acceptance of VAR analysis by making both the methodology and the data sets of volatility and correlation estimates for RiskMetrics publicly available on the Internet.

To illustrate the parametric approach we will calculate the VAR for one instrument—a Treasury bond futures contract.[3] In this example, we will estimate the VAR of a position consisting solely of a June 1996 Treasury bond futures contract purchased on May 24, 1996. The closing futures price for that day was 110. Since each Treasury bond futures contract is for the delivery of $100,000 in face value of bonds, each $1 change in the futures price results in a $1,000 change in the value of the position. VAR is usually estimated in terms of returns, rather than prices.

The return is calculated as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \cdot 100,$$

[3] A futures contract is an agreement to buy or sell an asset at a certain time in the future for a certain price. A Treasury bond futures contract is traded on the Chicago Board of Trade and is the most popular long-term interest rate derivative.

Table 1
*Daily Returns on a Bond Futures Contract*

| Date | Futures Price | Daily Return (%) |
|---|---|---|
| May 24, 1996 (Today) | 110.00 | |
| May 23, 1996 | 109.4063 | .542702 |
| May 22, 1996 | 110.1563 | −.68085 |
| May 21, 1996 | 109.5625 | .541928 |
| May 20, 1996 | 109.7813 | −.171086 |
| *** | *** | *** |
| May 31, 1995 | 111.25 | .674157 |
| May 30, 1995 | 111.125 | .112486 |

where $R$ is the daily return and $P$ is the price of the instrument. The daily returns on the bond futures contract are shown in Table 1.

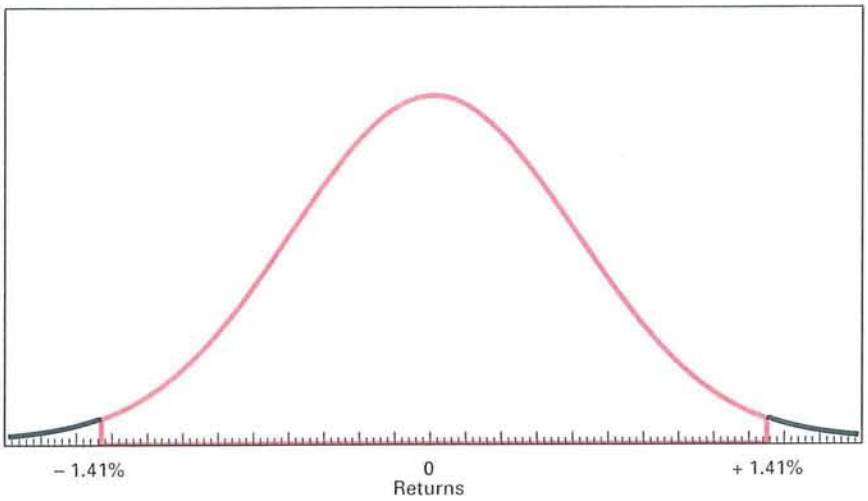
To calculate the one-day VAR of this position, we need to estimate the mean of the daily returns, and the volatility, as measured by the standard deviation. (Since the portfolio consists of only one instrument, we need not be concerned about correlations.) If these returns are governed by the normal distribution, then 95 percent of all returns will fall within 1.96 standard deviations of the mean return. Moreover, 98 percent of all returns will fall within 2.33 standard deviations of the mean return. The mean and the standard deviation in our case of Treasury future returns were found to be −0.00224 percent and 0.605074 percent, respectively. This means that 98 percent of all returns would fall between −1.41 percent and 1.41 percent and only 1 percent of returns will be lower than −1.41 percent (Figure 3).

To convert the negative return of 1.41 percent to a dollar amount, we recall that the futures price on May 24, the day for which we are calculating the VAR, was 110. From this we calculate a one-day VAR at the 1 percent probability level to be 1.41%/100 × 110 × $1,000 or $1,551.00. If the VAR estimate is correct, the daily loss on this position will exceed $1,551.00 no more than one day out of a hundred.

Suppose that the risk manager decides that a one-day holding period is too short, and that a one-week holding period is more appropriate. If, in addition to normality, we assume that returns are serially independent, meaning that a return on one day does not affect the return on any other day, then the standard deviation increases proportionately with the square root of time. Thus, if the one-day standard deviation of returns is 0.605074 percent, the standard

Figure 3

*VAR for Normally Distributed Futures Returns*




deviation for one week consisting of five trading days is $\sqrt{5} \times 0.605074$, or 1.3530 percent. This gives us a one-week VAR of $3,467.70 with a 1 percent probability level, which means that if we held the position for a week, we should not expect to lose more than $3,467.70 more often than in one week out of a hundred.

The two assumptions about the distribution of returns that underlie the parametric method—normality and serial independence—allow us to be very parsimonious in the use of data. Since volatilities and correlations are all we need to calculate the VAR at any confidence level for any holding period, it is unnecessary to have the historical returns themselves, which are used in the historical approach, as shown in Section VI.

## V. Does Volatility Change over Time?

In the preceding example, by calculating the volatility of the daily returns from a year of data, we implicitly assumed that the volatility of returns was constant throughout the year. However, volatility can change over time, sometimes quite abruptly, and it may make sense to pay more attention to the most recent observations in forecasting future volatility. For example, had we used only the last two months of returns, rather than a full year, we would have found the standard deviation to be 0.6513, rather than 0.605074, resulting in a VAR of $1,669.22, rather than $1,551.00. Clearly, the result would have been different still, had we used one month, six months, or five years of historical data on returns.

One popular way to estimate volatility is through exponential weighting of observations. This approach emphasizes more recent observations at the expense of the more distant ones because the weights assigned to past observations decline with time. The volatilities and correlations are updated every day in accordance with the most recent data, as the earliest observation is dropped from the historical series and the newest one is added.

The formula for the standard deviation ($\sigma$) of the daily return (R) and mean return $\mu$ with exponential weights based on a historical period of $N$ days is:

$$\sigma = \sqrt{(1 - \lambda)\sum_{i=1}^{N} \lambda^i (R_{N-i} - \mu)^2}.$$

The parameter $\lambda$ is known as the decay factor; it determines how fast the weight on past observations decays. The higher the $\lambda$, the slower is the rate of decay and the more weight is given to the more distant observations.[4] One study (Hendricks 1996) estimated volatilities for a number of decay factors and historical periods of different length for 1,000 simulated foreign exchange portfolios and found significant differences in the resulting volatilities.

[4] See J.P. Morgan, RiskMetrics Technical Document, 3rd ed. Chapter 2, Section 3 for details on the choice of decay factors. RiskMetrics currently uses a decay factor of 0.94 for all daily volatilities of the series it maintains (Peter Zangari, J.P. Morgan, personal communication).

Table 2
*Estimates of VAR Using Different Decay Factors*

| $\lambda$ = .94 (rapid decay) | | $\lambda$ = .97 (slower decay) | | $\lambda$ = .99 (slowest decay) | |
|---|---|---|---|---|---|
| $\delta$ | VAR | $\delta$ | VAR | $\delta$ | VAR |
| .5829 | $1,494.04 | .5503 | $1,410.48 | .4011 | $1,028.44 |

To check if our VAR estimate is sensitive to a choice of decay factors, we estimate the 1 percent probability level VAR for our bond futures returns using three different decay factors for a period of 50 days. Table 2 shows these estimates.

In the time period chosen for our example, the volatility of the return on the Treasury bond futures was increasing over time. Thus, the lower the weight placed on the more distant observations, the higher the estimate of volatility.

## *VI. The Historical Approach*

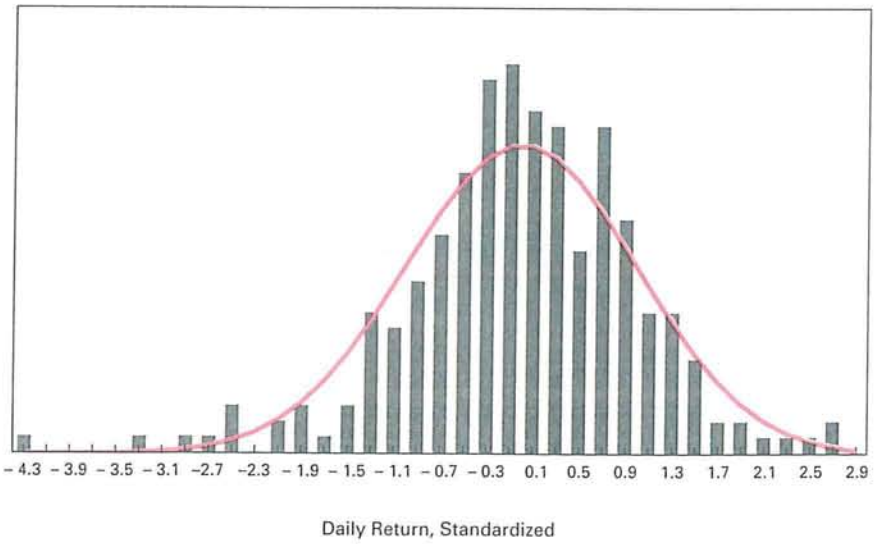
The simplicity and convenience of the normal distribution are powerful inducements for its use in VAR analysis, but this does not necessarily make its use appropriate. Since the early work of Mandelbrot (1963) and Fama (1965), most empirical research into the statistical properties of asset returns has found systematic deviations from normality. In particular, many studies have found that distributions of asset returns tend to exhibit kurtosis; namely, they are more peaked around the mean and have fatter tails than the normal distribution. Moreover, some, though not all, asset returns tend to be skewed to the left; that is, more unusually large negative returns are present than would have been expected if returns were normal.

To see if these observations apply to our chosen example of the bond futures contract, we construct a frequency distribution of the daily returns between May 1995 and May 1996. The resulting histogram is shown in Figure 4, with a normal distribution superimposed for comparison. The returns exhibit the typical pattern found in many asset returns: "fat tails" and left-skewness.

Fortunately, it is possible to calculate VAR without resorting to the assumption of normality, by using a simple historical method. This entails finding the lowest returns in the real historical data. To calculate VAR at the 1 percent probability level, we ranked the daily returns and identified the lowest 1 percent of returns. The first percentile return is 1.73 percent, giving us the daily VAR of 1.73%/100 × 110 × $1,000= $1,903.00, almost 23 percent greater

Figure 4

*Daily Returns on a U.S. Treasury Bond Future*



than the $1,551.00 calculated with the parametric approach.

If we want to recalculate the VAR for a different holding period without making the assumption of serial independence, we cannot simply multiply the daily VAR by the square root of time. Instead, we must recalculate all the returns for the new holding period, construct the new frequency distribution, and identify the appropriate percentiles.

The historical method has a number of advantages over the parametric method. First, it makes no explicit assumptions about the volatilities of returns and the covariances between them. Second, it makes no assumptions about the shape of the distributions themselves. In particular, it makes no assumption of normality. On the other hand, the historical approach lacks flexibility. Unlike the parametric method, it does not allow one to try different values for volatilities and correlations to test the sensitivity of VAR to these assumptions. In addition, it requires investors to obtain and maintain large amounts of actual, historical data. Long historical data series relevant for one's portfolio can be expensive or may not even exist. In contrast, the parametric approach requires just the parameters of distributions, if one is willing to delegate the estimation of those parameters to a third party.

*The historical method makes no explicit assumptions about the volatilities of returns and the covariances between them, or about the shape of the distributions themselves. On the other hand, it lacks flexibility and requires large amounts of actual, historical data.*

## *VII. Simulation Approaches—From Risk Factors to VAR*

Often, it is not appropriate to calculate VAR directly by estimating the probability distribution of returns on the instrument itself, as was done in the last example. If an institution has a large or complicated portfolio, it may be impossible or impractical to maintain historical data on all the instruments involved. Moreover, historical data do not exist for many instruments, particularly those that are customized. In those cases, the historical data set used to calculate the VAR will consist of returns not on the instruments themselves, but on their "risk factors," that is, other instruments or factors that influence their values. For example, for a domestic bond, the risk factor is the interest rate. For a bond denominated in foreign currency, the risk factors are the foreign interest rate and the exchange rate. For many equity derivatives, the main risk factor is the value of the S&P 500 index. For an S&P 500 option, the relevant risk factors are the value of the S&P 500 index, its volatility, the dividend yield on the index, and the risk-free interest rate.

In these cases, we can improve on the pure historical approach by using "historical simulation." Instead of looking at the volatility of the actual portfolio returns in the past, we will "simulate" the past portfolio returns by using the actual values of the risk factors and the *current* portfolio composition. Then, we can construct the empirical frequency distribution of the simulated portfolio returns by ranking them into percentiles and determining the VAR at the chosen confidence level.

### *Stochastic Simulation in Six Easy Steps*

Historical simulation shares one disadvantage with the simple historical approach: a lack of flexibility to investigate different assumptions. However, instead of using the past values of the risk factors, we can model these factors explicitly by specifying the underlying distributions and their parameters. Using these distributions and parameters, we can generate thousands of hypothetical scenarios for the risk factors and determine the portfolio value for each scenario. As in the historical simulation, the resulting portfolio returns can then be used to construct the empirical frequency distribution and determine the VAR at the desired confidence level. This is the approach generally known as the Monte Carlo, or "stochastic simulation."

As an illustration of this approach, we will calculate the one-day VAR for a position consisting of one call option, which gives the holder the right, but not the obligation, to buy an asset for a certain price. In this example, we will calculate the VAR for a call option on the S&P 500 index—a popular equity deriv-

ative traded on the Chicago Board of Trade. This particular option was bought on May 28, 1996 for $20.90 (the closing price for the contract on that day). The option expires on July 20, 1996 and gives the holder the right to buy the S&P 500 index for $670. (The actual value of the S&P 500 on May 28. 1996 was $674.9606.)

Finding the one-day VAR on May 28, 1996 involves the following six steps:

1. Calculate the daily returns on the S&P 500 and find the parameters of the normal distribution of returns.
2. Simulate returns on the S&P 500 for one day by generating random numbers from a normal distribution with the calculated parameters.
3. Use the simulated S&P 500 returns as input to calculate the simulated S&P 500 prices.
4. Use an option-pricing model to calculate the value of the call option at each simulated value of the S&P 500 index.
5. Calculate the one-day returns on holding the call option from the simulated call option prices.
6. Find the parameters of the distribution of call option returns and calculate the VAR.

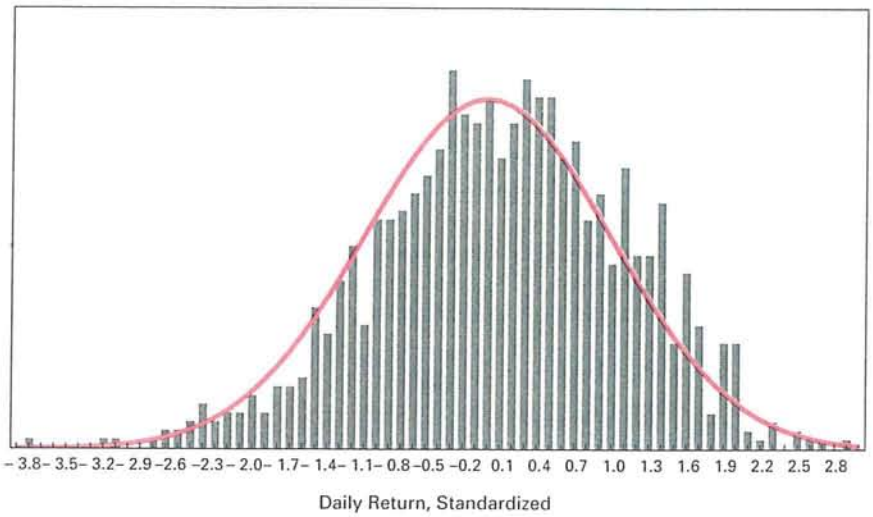
The crucial part of this process is the transformation from the distribution of the S&P 500 to the distribution of the option values. In this example, the option was valued using the standard Black-Scholes model modified to value a stock index.[5] The formula is as follows:

$$c = Se^{-q(T-t)}N(d_1) - Xe^{-r(T-t)}N(d_2),$$

where $d_1$ and $d_2$ are given by:

Figure 5

*Daily Returns on Simulated S&P 500 Call Premiums*




$$d_1 = \frac{\ln(S/X) + (r - q + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}}$$

$$d_2 = d_1 - \sigma\sqrt{T - t}.$$

The notation is as follows:

- $c$: the value of the call option;
- $S$: the price of the S&P 500 index;
- $T - t$: the time left until the expiration of the option (in this case, 53 days);
- $q$: the dividend yield of the S&P 500 (estimated to be 2 percent per year);
- $X$: the strike price of the option (in this case, $670);
- $N(x)$: the cumulative probability distribution for a standardized normal variable (that is, the probability that such a variable will be less than $x$);
- $r$: risk-free rate of interest, in this case, the federal funds rate, or 5.5 percent per year;
- $\sigma$: volatility of the S&P 500 index.

The distribution of the resulting simulated returns is depicted in Figure 5. To calculate VAR, we rank the returns into percentiles and identify the 1 percentile return. This happens to be −0.3560. The price of the call option on May 28, 1996 was 20.90.

[5] The model for valuing options for non-dividend-paying stocks was developed by Black and Scholes (1973). The formula for valuing options on stock indexes paying dividends was derived by Merton (1973).

Every option contract on the S&P 500 is for $100 of the index. Thus, the dollar value of VAR for one contract is 0.3560 × 20.90 × $100 = $744.04. This means that the chance is 1 in 100 that we will lose more than $744.04 on the call option contract in one day.

It should be noted that, in this example, volatility was calculated from one year of data with all observations weighted equally. Of course, a longer or shorter historical period could be used along with various exponential weighting schemes to generate a different estimate of volatility of the S&P 500 index, which would result in a different VAR for the call option.

## *VIII. Conclusion*

In many financial circles, the reputation of value at risk stands as high as that of motherhood and apple pie. But, as with motherhood and apple pie, a good concept is not enough—good implementation is equally important. Value-at-risk analysis is a general framework that covers models with a wide variety of assumptions and methods of calculation, and, inevitably, it produces a wide variety of results. Our examples demonstrate that varying assumptions about distributions and methods of calculating volatilities produced quite different estimates.

Value at risk is a useful tool for risk management but it is not a panacea. One limitation is that it focuses on a single arbitrary point on the distribution of profits and losses, while it would be more useful to have a representation of the whole distribution. A second limitation of value at risk is that it tells us little about how risks are to be measured in extreme market conditions. During market crises, correlations between asset prices break down, liquidity disappears, and price data may not be available at all. Modeling risk under such conditions would require some sense of the concentration of ownership of different securities, information that most market participants would be reluctant to disclose, for competitive reasons.

Value at risk is often put forward as the way to aggregate risks across the whole institution. While integrating disparate risks is the ultimate purpose of global risk management, the use of value at risk for this purpose is problematic. Value-at-risk analysis works best for frequently traded instruments for which market values are easily available. Value at risk first took hold at the derivatives desks in the trading rooms of a few large banks, because they had both the expertise and the need to estimate and aggregate the market risks of many dissimilar instruments. From derivatives desks it spread to other trading desks, such as those for bonds and currency, and it is now beginning to be applied beyond trading to the broader arena of asset–liability management. However, many bank assets and liabilities, in particular deposits and loans, have long-term horizons and are not actively traded. Thus, they have poor or nonexistent price data and marking them to market on a day-to-day basis would be both impractical and misleading. Value at risk has the same limitations for life insurance companies, because these institutions have long-term, nontraded liabilities. Other methods may prove superior to value at risk as a global measure of risk. One alternative is a measure of how much one is willing to pay to eliminate risk, or the price of purchasing a guarantee to avoid a loss of a certain magnitude. Value at risk, which focuses on the distribution of possible losses, is only one element in the valuation of such a guarantee.

Overall, value at risk constitutes a useful though limited family of techniques for measuring risk. It is most useful in measuring short-term risk of traded instruments in normal market conditions. An additional benefit is that its use has created a common language for discussions about risk, and it has prompted more dialogue about risk issues. However, successful risk management is a much broader task, which depends crucially on appropriate incentives and internal controls.

## *References*


Black, Fischer and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, vol. 81 (May-June), pp. 637–54.

Fama, Eugene. 1965. "The Behavior of Stock Market Prices." *Journal of Business*, vol. 38, pp. 34–105.

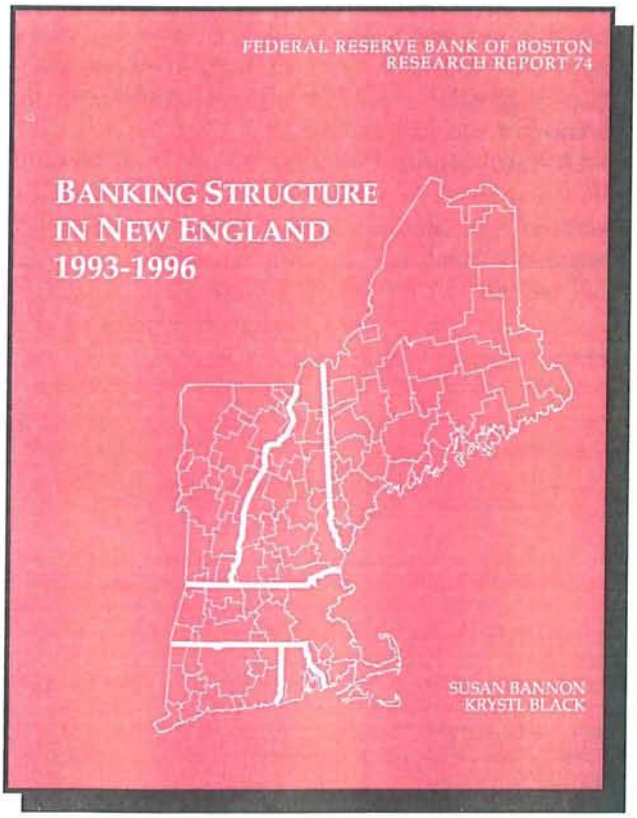
Group of Thirty. 1993. *Derivatives: Practices and Principles*. Global Derivatives Study Group. Washington, DC.

Hendricks, Darryl. 1996. "Evaluation of Value at Risk Models Using Historical Data." Federal Reserve Bank of New York *Economic Policy Review*, April, pp. 39–70.

Mandelbrot, Benoit. 1963. "The Variation of Certain Speculative Prices." *Journal of Business*, vol. 36, pp. 394–419.

Merton, Robert. 1973. "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, vol. 4 (Spring), pp. 141–83.

## Research Report




Research Report No. 74 has just been issued. *Banking Structure in New England 1993-1996*, by Susan Bannon and Krystl Black, summarizes structural changes in New England's banking industry that occurred between June 30, 1993, and April 30, 1996, including mergers and acquisitions, bank holding company formations, bank openings and closings, and name changes. An update of Research Report No. 73, the new report uses 1995 data to rank New England commercial banking and thrift organizations by total consolidated New England deposits. Rankings are presented for all of New England, for each state, and within local banking markets. For each local banking market, Herfindahl-Hirschman Indexes and three-firm concentration ratios are provided. The report begins with a discussion of relevant antitrust issues.

Research Report No. 74 is available without charge. Requests should be sent to Research Library - D, Federal Reserve Bank of Boston, P.O. Box 2076, Boston, MA 02106-2076. The fax number is (617) 973-4292, and the e-mail address is boston.library@bos.frb.org.

# The Saving Mystery, or Where Did the Money Go?

*Lynn Elaine Browne with Joshua Gleason*

*Senior Vice President and Director of Research, and Senior Research Assistant, respectively, Federal Reserve Bank of Boston.*

Of great concern and puzzlement to many has been the decline in the U.S. personal saving rate. From 8 percent of personal income 20 years ago, saving has fallen to less than 4 percent. This is a matter of concern because saving and investment are closely linked, and investment is believed critical to productivity gains and a rising standard of living. In addition, the decline in saving is sometimes presented as a national character flaw and evidence of a more profligate and self-centered population. The decline in saving is also a source of puzzlement because it runs counter to many people's perception of what is happening. In particular, the large inflows into mutual funds in recent years and the strong performance of the stock market seem inconsistent with statistics showing that the saving rate has fallen to unprecedented lows.

This article will investigate the decline in saving, focusing on "where the money went." It will look at both the nature of the consumption that the decline in saving has supported and the changes in households' assets and liabilities that have accompanied this decline. It will show how households' increasing investment in mutual funds can coexist with a dwindling saving rate. And while it may not alleviate concerns that our saving is too low, it will argue that the real issue is not saving per se but how to boost productivity growth and raise standards of living.

## *Declining Saving*

The basic "problem" appears in Figure 1. Total saving has fallen from more than 20 percent of GDP in the 1960s, to 20 percent in the 1970s and early 1980s, to only 15 percent in the 1990s. Paralleling the falloff in saving has been a decline in investment, with private domestic investment falling especially sharply since the mid 1980s. (See Table 1 for investment's share of GDP and the components of saving.)

An important culprit here is the federal government. During the

Table 1
*Saving and Investment as a Percentage of GDP*

| | Total Gross Saving | Personal Saving | Gross Business Saving | Gross Government Saving | Federal Government Saving | Gross Investment | Gross Private Domestic Investment |
|---|---|---|---|---|---|---|---|
| 1959 | 21.5 | 4.8 | 11.5 | 5.2 | 2.5 | 21.1 | 15.5 |
| 1960 | 21.6 | 4.4 | 11.2 | 6.0 | 3.4 | 20.9 | 15.0 |
| 1961 | 21.4 | 5.2 | 11.1 | 5.2 | 2.5 | 20.8 | 14.3 |
| 1962 | 21.8 | 5.0 | 11.5 | 5.2 | 2.4 | 21.4 | 15.0 |
| 1963 | 21.9 | 4.6 | 11.6 | 5.7 | 2.8 | 21.4 | 15.1 |
| 1964 | 22.0 | 5.3 | 11.7 | 5.0 | 2.0 | 21.7 | 15.3 |
| 1965 | 22.4 | 5.3 | 12.0 | 5.1 | 2.2 | 22.2 | 16.4 |
| 1966 | 21.8 | 5.0 | 11.9 | 5.0 | 2.0 | 22.1 | 16.6 |
| 1967 | 20.9 | 5.9 | 11.5 | 3.6 | .7 | 21.0 | 15.4 |
| 1968 | 20.4 | 5.1 | 10.9 | 4.4 | 1.3 | 20.4 | 15.4 |
| 1969 | 20.7 | 4.8 | 10.4 | 5.5 | 2.5 | 20.4 | 15.8 |
| 1970 | 19.1 | 5.9 | 10.0 | 3.1 | .2 | 19.2 | 14.5 |
| 1971 | 19.1 | 6.1 | 10.9 | 2.1 | −.8 | 19.6 | 15.6 |
| 1972 | 19.8 | 5.1 | 11.2 | 3.4 | −.2 | 20.1 | 16.6 |
| 1973 | 21.5 | 6.5 | 11.0 | 4.0 | .6 | 21.7 | 17.6 |
| 1974 | 20.2 | 6.5 | 10.4 | 3.4 | .3 | 20.5 | 16.4 |
| 1975 | 18.3 | 6.4 | 12.1 | −.2 | −3.1 | 19.0 | 13.8 |
| 1976 | 18.7 | 5.3 | 12.2 | 1.3 | −1.8 | 19.8 | 15.8 |
| 1977 | 19.5 | 4.6 | 12.7 | 2.3 | −1.0 | 20.4 | 17.6 |
| 1978 | 20.8 | 4.9 | 12.8 | 3.2 | −.1 | 21.6 | 18.8 |
| 1979 | 21.1 | 5.1 | 12.5 | 3.5 | .5 | 22.2 | 18.8 |
| 1980 | 19.7 | 5.8 | 11.8 | 2.0 | −1.0 | 20.6 | 16.7 |
| 1981 | 20.9 | 6.4 | 12.3 | 2.2 | −.7 | 21.4 | 17.9 |
| 1982 | 18.7 | 6.3 | 12.5 | −.2 | −2.9 | 18.6 | 15.5 |
| 1983 | 16.8 | 4.8 | 12.9 | −.8 | −3.8 | 17.8 | 15.6 |
| 1984 | 19.3 | 6.0 | 12.9 | .4 | −2.9 | 19.4 | 18.3 |
| 1985 | 17.9 | 4.9 | 12.6 | .4 | −2.8 | 17.9 | 17.1 |
| 1986 | 16.3 | 4.4 | 11.6 | .2 | −2.9 | 16.8 | 16.3 |
| 1987 | 16.6 | 3.6 | 11.9 | 1.1 | −1.6 | 16.3 | 15.9 |
| 1988 | 17.4 | 3.7 | 12.3 | 1.4 | −1.3 | 16.4 | 15.3 |
| 1989 | 16.7 | 3.5 | 11.5 | 1.7 | −1.0 | 16.9 | 15.2 |
| 1990 | 15.7 | 3.6 | 11.4 | .7 | −1.6 | 16.0 | 13.9 |
| 1991 | 15.8 | 4.2 | 11.6 | .1 | −2.2 | 16.0 | 12.4 |
| 1992 | 14.5 | 4.4 | 11.2 | −1.1 | −3.4 | 15.2 | 12.7 |
| 1993 | 14.3 | 3.3 | 11.4 | −.4 | −2.8 | 15.2 | 13.3 |
| 1994 | 15.2 | 2.8 | 11.7 | .7 | −1.7 | 15.7 | 14.6 |
| 1995 | 15.8 | 3.3 | 11.3 | 1.1 | −1.2 | 15.8 | 14.7 |

Source: U.S. Bureau of Economic Analysis.

1960s, the federal government ran a small surplus, and thus appears as a net contributor to saving.[1] In the 1970s, however, the federal government began to incur budget deficits; these deficits became chronic in the 1980s and 1990s.

Initially, the larger federal deficits were offset by higher personal and business saving. In the mid 1980s, however, personal saving fell sharply. This can be seen more clearly in Figure 2, which shows personal saving relative to both personal income and after-tax or disposable income. (Personal income and saving include the income and saving of nonprofit institutions serving individuals as well as the income and saving of individuals.)

The decline in personal saving becomes even more critical when one considers that the large federal deficits of the 1980s and 1990s are primarily the result of increased transfer payments.[2] Transfers are additions to personal income for which the recipient performs no current service. They include Social Security benefits, payments made under government-provided health insurance, unemployment insurance benefits, and payments to federal retirees. At the state and local level, expenditures on medical care for the needy and other public assistance payments, as well as retiree benefits, are the dominant transfers.

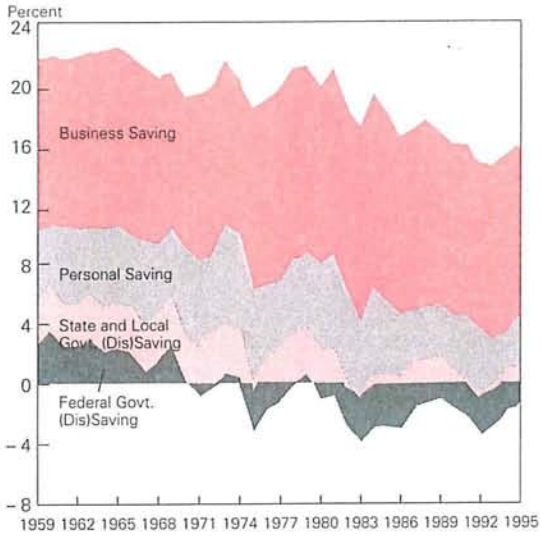
In large part because of the rapid growth in transfers, personal income has increased relative to GDP; but since the mid 1980s, a smaller fraction of personal income has been saved. Thus, government

[1] The federal government contributed to saving in the 1960s by financing its own investment expenditures out of taxes and other receipts. Expenditures on defense equipment account for the bulk of federal government investment.

[2] Increasing expenditures on defense in the early 1980s contributed to the emergence of large federal deficits, but defense spending was subsequently curtailed; and the federal government's direct expenditure on goods and services accounted for a smaller fraction of GDP in the mid 1990s than 10 years earlier.

Figure 1

*Saving as a Share of GDP*




Source: U.S. Bureau of Economic Analysis.

dissaving has, in effect, financed personal consumption. This is precisely the opposite of what should happen, according to some economists, who argue that government dissaving will elicit *increased* private saving, as households will perceive and offset the burden that government deficits impose on future generations.

Government transfer payments also increased very rapidly in the 1960s and 1970s, but consumption's share of GDP did not rise, as the fraction of income going to both taxes and saving increased. In contrast, consumption rose from 62 percent of GDP in the late 1970s to 65 percent in the mid 1980s to 68 percent in 1995, as federal dissaving augmented personal income, even as the personal saving rate declined.

Business saving has contributed only modestly to the decline in total saving but its composition has shifted, with retained earnings accounting for a smaller fraction of business saving and depreciation charges a larger share.[3] Shorter-lived equipment explains some of the shift; in addition, corporations have been paying out a larger fraction of profits as dividends in the 1980s and 1990s than they did previously. As in the case of government transfers, the increase in dividend payments represents an increase in personal income.

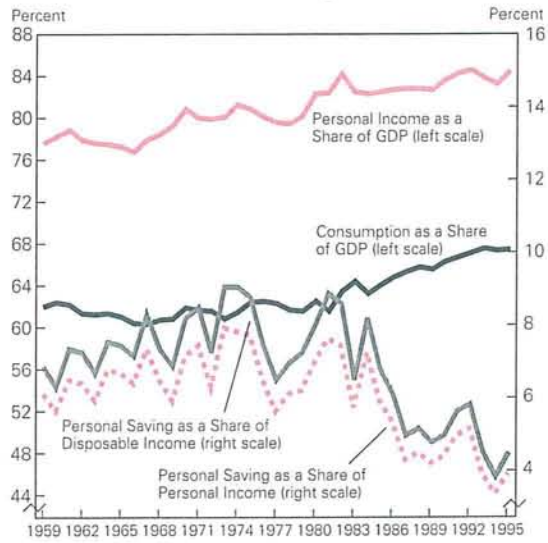
The key point, however, is that federal government dissaving and a decline in the personal saving rate are the reasons that saving has fallen relative to GDP. Moreover, federal dissaving has taken the form of increases to personal income, which have then been consumed. The result has been a sizable increase in the fraction of output going to personal consumption. The following section examines the changes in consumption patterns that have accompanied these shifts.

## *Where the Saving Went—The Consumption Story*

The allocation of personal income among major uses appears in Figure 3. The chart draws attention to a point that should be borne in mind in any discussion of saving: While the share of income going to saving has fallen by half over the past 10 years, this change is fairly small relative to the shifts among expenditure

Figure 2

*More Income, More Consumption, Less Saving*




Source: U.S. Bureau of Economic Analysis.

[3] Depreciation charges, called capital consumption allowances (CCA) in the National Income and Product Accounts (NIPA), represent the bulk of business savings. Charges for accidental damage are also included in CCA.

## What's Happened to Investment?

A key reason for concern over the declining saving rate is that it implies lower investment. But while gross saving and investment fell from more than 21 percent of GDP in the 1960s to under 16 percent in the 1990s, less than a third of the decline is due to a falloff in the domestic investment of the private sector. Much of the drop is attributable to government investment and to the United States' shifting from being a net investor abroad to a net borrower.

At the federal level, investment in defense equipment is a significantly smaller fraction of GDP now than it was in the 1960s, with the falloff especially pronounced since the late 1980s. State and local governments' investment in public buildings is the other area of notable decline.

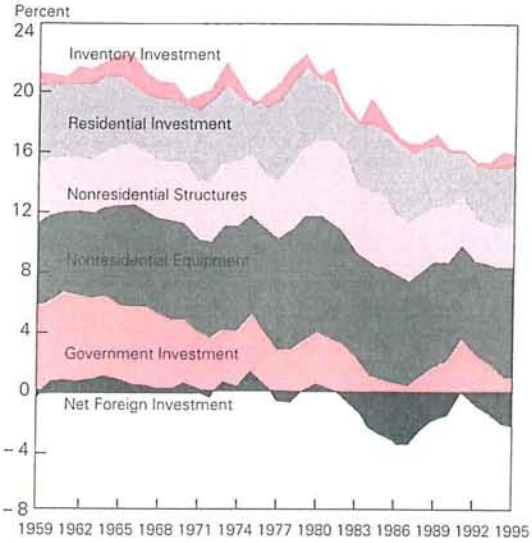
In addition, the United States has offset a portion of the decrease in its own saving by drawing upon the saving of the rest of the world. During the 1960s and 1970s, the United States was a net investor in the rest of the world, but in the 1980s the United States became a net borrower (Figure B1). Tapping foreign saving has helped sustain private investment in this country, presumably with beneficial effects for U.S. productivity and wage growth. However, reliance on foreign rather than domestic saving means that a portion of the nation's productive capacity and income generation will be devoted to supporting foreign rather than domestic consumption in the years ahead.

Focusing on private domestic investment, the shares of GDP invested in nonresidential structures and housing have fallen. The reduction in nonresidential structures is especially pronounced in comparison with the early 1980s, which were characterized by a boom in office building. Investment in business equipment has not fallen and is actually a larger fraction of GDP today than it was in the 1960s.

Because equipment is shorter-lived than structures and because the composition of equipment has shifted towards short-lived information processing equipment, depreciation charges (called Capital Consumption Allowances in the National Income and Product Accounts) have increased relative to gross investment (Figure B2). Thus, the falloff in investment's share of output appears even steeper when depreciation is netted out. Nevertheless, since the prospect of declining investment in business equipment is seen by some as the most pernicious consequence of declining saving, the continued strength here should be some comfort.

Figure B1

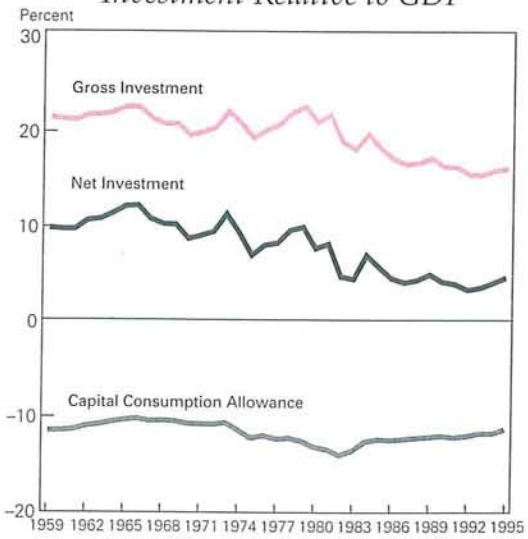
*Gross Investment Relative to GDP*




Source: U.S. Bureau of Economic Analysis.

Figure B2

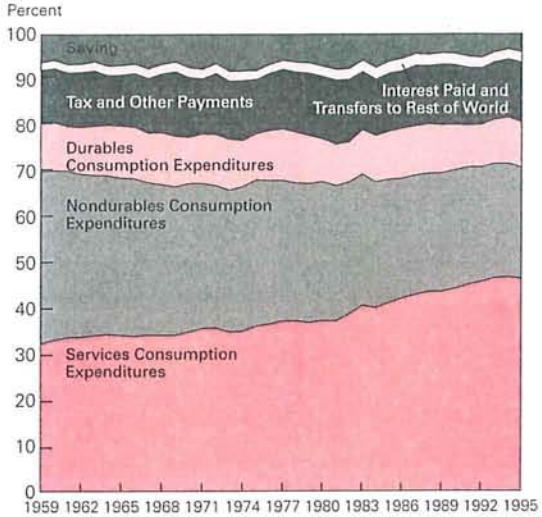
*Gross Investment, Depreciation, and Net Investment Relative to GDP*




Source: U.S. Bureau of Economic Analysis.

Figure 3

*The Disposition of Personal Income*




Source: U.S. Bureau of Economic Analysis.

Table 2

*Major Expenditure Categories as a Percent of Personal Income*

Percent

| Selected Years | Durable Goods | Nondurable Goods | Services |
|---|---|---|---|
| 1959 | 10.8 | 37.7 | 32.3 |
| 1965 | 11.4 | 34.5 | 34.1 |
| 1975 | 10.1 | 32.0 | 36.1 |
| 1985 | 10.5 | 27.0 | 41.2 |
| 1995 | 9.9 | 24.3 | 46.4 |
| Annual rate of change in expenditures, 1959–1995 | | | |
| Expenditures | 7.6 | 6.6 | 9.0 |
| Price | 2.6 | 4.1 | 5.1 |
| Quantity | 4.9 | 2.4 | 3.7 |
| Annual rate of change in expenditures, 1985–1995 | | | |
| Expenditures | 5.3 | 4.8 | 7.2 |
| Price | 1.7 | 2.9 | 4.4 |
| Quantity | 3.5 | 1.9 | 2.7 |

Source: U.S. Bureau of Economic Analysis.

categories. Moreover, because saving is small and because it is calculated as a residual—what is left over from personal income after taxes, consumption, and interest payments have been deducted—it is sensitive to errors in the measurement of personal income and consumption.[4] Thus, the precise magnitude of the decline in the personal saving rate is less the issue than the rising fraction of income and output going to consumption.

The patterns shown in Figure 3 are based on current dollars and measure the fraction of income absorbed by different categories of expenditure. This is the appropriate concept for the purpose at hand, identifying those consumption areas where rising expenditures have squeezed out saving. As can be seen in the lower panel of Table 2, however, differences in the rates at which prices have increased cause the patterns of consumption shown by current dollar expenditures to vary considerably from those shown by the inflation-adjusted quantity estimates that are more commonly used in economic analysis.

Services plays the critical role. The fraction of income absorbed by services consumption has risen from a third in the 1960s and 1970s to over 40 percent in the mid 1980s to more than 45 percent today. This increase has been driven primarily by the rapid growth in the prices of services, although distinguishing between price and quantity changes is not very meaningful for many services.[5] In contrast to services, the share of income allocated to durable goods has remained roughly constant at 10 percent, with relatively rapid increases in the quantity of durable goods consumed offset by relatively small price increases. Most of the increase in services' share of income has come out of the share going to nondurable goods, where the quantities consumed, especially of food products, have grown slowly. The balance has come out of saving.

Digging deeper, the growth in services has been driven primarily by increased expenditures on medical care and "other" services (Table 3). The share of

[4] Saving also includes an imputed element, of which the largest component is net purchases of owner-occupied housing, less the consumption of fixed capital on housing. Purchases of housing are treated as saving (and investment) because housing is a long-lived good and delivers services over many years. The value of these housing services, however, is treated as part of consumption and is estimated as the rent that owners would pay to occupy their housing.

[5] Many economists would contend that distinguishing between price and quantity increases is not meaningful for services because defining the appropriate unit of output is very difficult and because, for such important services as medical care, quality changes have been profound.

Table 3
*Major Services Expenditures Relative to Personal Income*
Percent

| Selected Years | Total Services | Housing Services | Household Operation | Transport Services | Medical Services | Other Services |
|---|---|---|---|---|---|---|
| 1959 | 32.3 | 11.4 | 4.7 | 2.7 | 4.2 | 9.2 |
| 1965 | 34.1 | 11.8 | 4.8 | 2.6 | 5.0 | 9.9 |
| 1975 | 36.1 | 11.2 | 4.9 | 2.7 | 7.1 | 10.3 |
| 1985 | 41.2 | 11.8 | 5.3 | 2.7 | 9.4 | 11.8 |
| 1995 | 46.4 | 12.2 | 4.8 | 3.2 | 12.9 | 13.4 |

*Significant "Other" Services*
Percent

| Selected Years | Personal Care | Personal Business | Recreation | Education and Research | Religious and Welfare | Expenditure in U.S. by Foreigners | Foreign Travel |
|---|---|---|---|---|---|---|---|
| 1959 | 1.6 | 3.5 | 1.6 | 1.0 | 1.3 | −.3 | .5 |
| 1965 | 1.5 | 3.8 | 1.7 | 1.2 | 1.4 | −.3 | .6 |
| 1975 | 1.0 | 4.0 | 1.9 | 1.6 | 1.5 | −.4 | .7 |
| 1985 | .8 | 5.3 | 2.2 | 1.5 | 1.8 | −.7 | .8 |
| 1995 | .8 | 6.1 | 2.7 | 1.8 | 2.3 | −1.2 | .8 |

Source: U.S. Bureau of Economic Analysis.

income devoted to medical care alone has grown from 9 percent in the mid 1980s to almost 13 percent in the mid 1990s. Within medical care, increases in expenditures have been rapid in all major categories. Hospital services, the largest component of medical care, accounted for about one-third of the increase, although expenditures on physicians and other professional services providers (such as chiropractors and private nurses) have grown more rapidly over the past 10 years.[6]

Of the growing share of income allocated to "other" services, expenditures on personal business account for roughly half. Within personal business, the largest component is "services provided without payment by financial intermediaries." This is an imputed element and also appears as an addition to the interest portion of personal income.[7] It grew very rapidly in the late 1970s and early 1980s. More recently, the fastest-growing components have been expenditures for brokerage fees and investment counseling and bank service charges. Also in the personal business category are life insurance and legal expenses.

Expenditures on recreation services, education, and religious and welfare activities make up most of the balance of "other" services. Education includes the expenditures of private educational institutions, nursery through university, and tuition paid to government institutions. Religious and welfare activities include the expenditures of nonprofit institutions, such as churches, social welfare organizations, museums, and foundations, and fees paid to proprietary and government institutions. Day care falls in this category. Individually each of these elements is very small; but altogether, "other" services' share of income has risen by about as much as saving's share has fallen.

The other large element of services consumption, housing services, has grown only slightly faster than personal income. The largest component of housing services is not a market transaction, but an estimate of what homeowners would pay if they rented rather than owned their dwellings. Homeowners, as landlords, are then credited with income from the "rental" of their properties, after deducting depreciation and certain other costs.[9] Net purchases of owner-occupied housing are viewed as investment and thus, are part of saving.

[6] Expenditures on medical services are broken down according to the entity providing the service—physicians, dentists, other professional services, hospitals and nursing homes, and insurance. The insurance component consists of premiums paid for health insurance less benefits and dividends. Thus, it represents payment to insurers for their insurance services rather than the medical services that are financed by insurance benefits. The latter appear as expenditures on various medical care providers.

[7] "Services provided without payment by financial intermediaries" appears as part of consumption and also as part of the interest component of personal income. In effect, the national accounting framework assumes that financial intermediaries impose charges for certain services for which there are no explicit fees, and that these charges are offset by interest payments to depositors.

[8] Expenditures on education services do not include expenditures for the meals, rooms, and other non-educational services that these institutions provide. These are counted in the relevant categories of consumption.

[9] The rationale behind this approach is that the national income and product accounts should be neutral with respect to the ownership of residential property. Estimates of economic output should not be affected by whether people rent housing or own it directly. The imputed income from housing appears in the net interest and rental income of personal income.

The housing imputations are large compared to personal saving and Kopcke, Munnell, and Cook (1991) have argued that the NIPA

The point of this rather detailed rundown of how households allocate their incomes is that it does not square well with characterizations of low saving rates as resulting from the selfishness and profligacy of today's generations as compared with their predecessors. The growth in consumption has been driven primarily by the growth in medical expenditures and by certain personal business and other services expenditures that do not fit most people's idea of consumption. Indeed, expenditures on brokerage and investment counseling and life insurance expenses are costs of saving. In addition, some of the expenditures on education and welfare activities, as well as some health care expenditures, could be regarded as investments, in that they augment human capital and, thus, enhance the economy's productive capacity.

*Where did the saving go? To support the consumption of services, medical services primarily but also personal business, education, religious activities, and recreation. Only the last fits the customary image of consumption.*

Some of these expenditures are also unusual in that the level of expenditure is not really controlled by individual consumers. Thus, consumption of medical services includes not only consumers' out-of-pocket costs but also the expenditures made on their behalf by their private health insurance plans and by the government through Medicare and medical public assistance. Thus, there is a disconnect between the consumption of medical care and the purchase decision. While people can influence the level of medical care to which they have access through their choice of insurance coverage and as voters and taxpayers, once they have purchased insurance or are qualified for government assistance, their consumption of medical services is not constrained by their income.

Because government-financed medical services are counted as transfer payments and augment personal income as well as consumption, their rapid growth (from virtually nothing in the 1960s to 5 percent of personal income in the mid 1990s) does not provide a direct explanation for the decline in the personal saving rate, although it has contributed to government dissaving and to the increase in consumption relative to GDP.[10] In contrast, the growth in employers' contributions to medical insurance represents a shift in the composition of personal income rather than an addition. Had employers' contributions not increased, wage and salary disbursements presumably would have grown faster. But while employees can choose to consume or save the income they receive as wages and salaries, they can only consume employers' contributions to health insurance. Such contributions now amount to about 8 percent of wages and salaries, compared to 5 percent in the mid 1980s and 2 percent in the mid 1960s.

Thus, the answer to the question "where did the saving go?" is that it went to support the consumption of services, medical services primarily but also personal business, education, religious and welfare activities, and recreation. Only the last fits the customary image of consumption.

But whether consumption in the normal sense or not, these activities have absorbed income that might otherwise have been saved. Yet to many people, the situation appears very different. The popularity of 401(k) pension plans and large inflows into mutual funds, plus the records set by the stock market, as well as numerous stories in the financial press about baby-boomers having gotten the saving religion, all tell them that the personal saving rate should be rising.

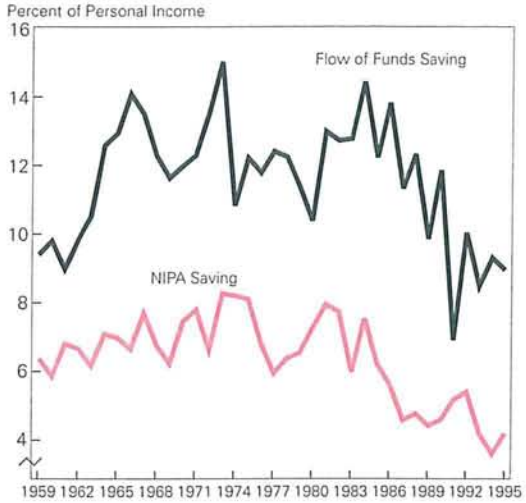
As noted above, these saving figures, from the National Income and Product Accounts (NIPA), are calculated as a residual and show what is left over from income after various outlays are deducted. Another measure of saving, calculated by adding up households' accumulation of assets and liabilities, addresses more directly the phenomena that seem to argue for a higher saving rate. But this measure, from the Federal Reserve Board's Flow of Funds (FOF) estimates, also shows a sharp decline since the mid 1980s.

treatment of owner-occupied housing results in an understatement of saving.

[10] In a detailed analysis of spending by age cohort, Gokhale, Kotlikoff, and Sabelhaus (1996) conclude that government redistribution from younger to older generations and an increase in the propensity of older people to consume, especially in the form of medical care, are responsible for the decline in the U.S. saving rate.

Figure 4

*Flow of Funds Measure of Saving versus Saving in the National Income and Product Accounts*




Source: Board of Governors of the Federal Reserve System, Flow of Funds data base; U.S. Bureau of Economic Analysis.

## *Where the Saving Went—The Asset Story*

The FOF and NIPA measures of saving are compared in Figure 4. The gap between the two measures is due largely to the treatment of households' purchases of motor vehicles and other consumer durables. Both saving measures treat net purchases of housing as an investment, from which the depreciation of the existing stock of housing is deducted. FOF saving also treats purchases of consumer durables this way, on the grounds that these are long-lived goods and deliver transportation, entertainment, and other services over many years. Increases in consumer credit used to finance durables purchases must be subtracted to determine the net contribution to saving.[11]

FOF saving is, thus, the sum of net acquisitions of financial assets and net investment in tangible assets less net increases in financial liabilities. As can be seen from Figure 5, however, total saving has tracked the acquisition of financial assets for the past 20 years, and the main reason for the decline in saving since the mid 1980s has been a reduction in the accumulation of financial assets. One asset accounts for all of this falloff—money.

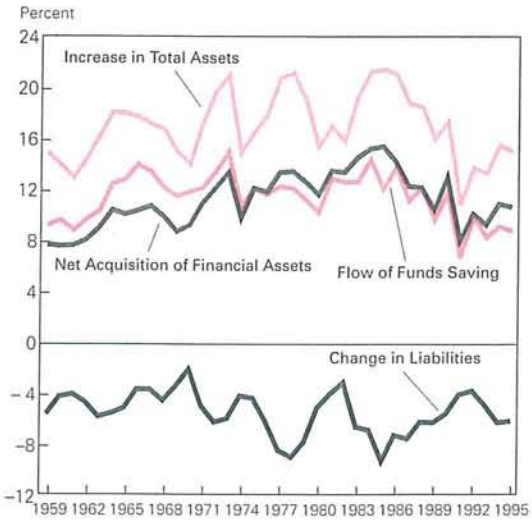
The composition of saving through the acquisition of financial assets appears in Figure 6. In the 1970s, households and nonprofit institutions saved roughly 7 percent of personal income in the form of various types of deposit accounts and currency. Beginning in the mid 1980s, the share of personal income saved as deposits fell sharply and in the early 1990s, almost no saving took this form. (Time and savings deposits account for most of this decline.)

Some of the funds that were previously saved as deposits may have gone into mutual funds, which emerged as an important savings vehicle in the second half of the 1980s. The increase in mutual funds was not enough to offset the falloff in money saving, however. Moreover, a substantial portion of the inflow into mutual funds over the past 10 years has been offset by sales of corporate equities. For many savers, mutual funds may not be so much an addition to saving as an alternative to direct ownership of corporate securities.

Pension fund reserves are the most important way in which financial saving currently takes place. These include the assets of both defined benefit plans

Figure 5

*Change in Assets and Liabilities Relative to Personal Income*

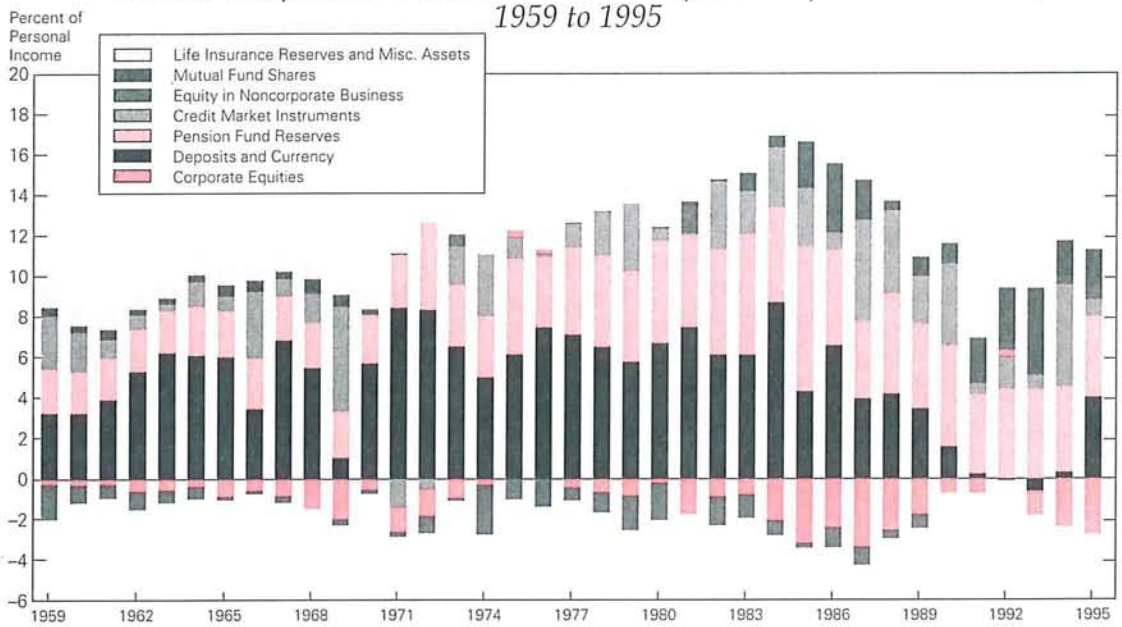



Source: Board of Governors of the Federal Reserve System, Flow of Funds data base.

[11] Of course, not all consumer credit is used to finance consumer durables and some mortgage debt is not really motivated by the decision to purchase a home. Thus, it is not strictly accurate to set all of consumer and mortgage debt against the acquisition of tangible assets; a portion represents general liabilities.

Figure 6

*Selected Components in Households' Net Acquisition of Financial Assets, 1959 to 1995*




Source: Board of Governors of the Federal Reserve System, Flow of Funds data base.

and defined contribution plans, such as 401(k) plans. Saving through pensions represents about 4 percent of personal income. This is less than in the first half of the 1980s, but more than in the 1960s and 1970s. Since the rules clarifying the legality of 401(k) plans were issued in 1981, at least some of the bulge in pension reserves in the early 1980s is presumably a response to that event.

Saving can also occur through the acquisition of tangible assets. Since the mid 1970s, however, increases in households' mortgages and consumer credit liabilities have offset net investment in housing and consumer durables (new expenditures less depreciation of the existing stock of housing and durables). Thus, by this calculation, tangible assets have made virtually no contribution to saving for the past 20 years. (Again, see footnote 11.)

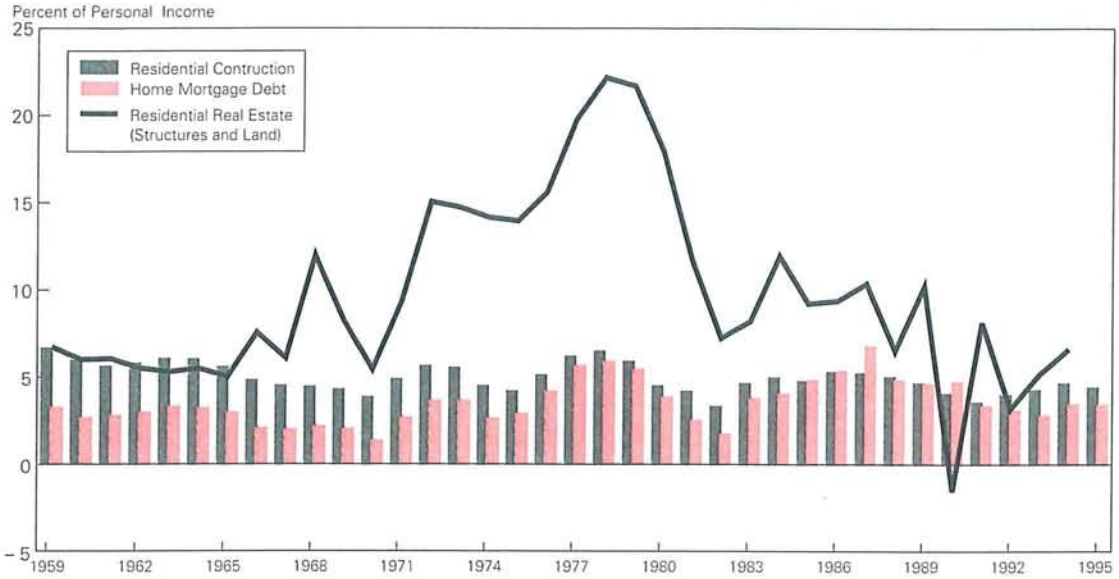
Back in the 1960s, in contrast, households' net investment in housing and consumer durables exceeded additions to their mortgage and consumer credit liabilities by 2 to 3 percent of personal income, in some years augmenting saving through financial assets by a third. As can be seen in Figure 7, households began to take on more mortgage debt in the 1970s, most likely in response to the "hot" housing market of the time. Prices rose rapidly, requiring prospective home-buyers to borrow more, even as the higher value of the underlying collateral made lenders more willing to lend.[12] Although the rapid escalation in home prices slowed in the mid 1980s, households continued to add to mortgage debt at the previous rate, as the introduction of home equity loans and increased competition arising from securitization and the growing importance of mortgage banks made mortgages more accessible.

In summary, the FOF data show that the decline in saving over the past 10 years has been concentrated in bank deposits. Although inflows into mutual funds have been substantial, they have not been sufficient to offset the falloff in deposit saving. In addition, sales of corporate equities have negated a portion of the mutual fund inflow. On the liabilities side, increases in mortgage debt have offset additions to saving from new home construction.

[12] The late 1970s were characterized by very high turnover in existing homes relative to new construction. Thus, the number of transactions involving mortgages increased relative to the number representing new saving.

Figure 7

*Change in the Value of Residential Real Estate versus Growth in Residential Construction and Mortgage Debt*




Source: Board of Governors of the Federal Reserve System, Flow of Funds data base.

## *Interpretation and Policy Discussion*

These shifts among assets, as well as the expenditure patterns discussed in the preceding section, raise interesting questions of interpretation. They highlight a dilemma facing those concerned about low saving rates: The individual's perception of the need for saving may differ from society's. Economists commonly approach the question of why people save using a life-cycle model. According to this model, people's consumption, and therefore saving, is based on the present value of the resources they expect to be available to them over the remainder of their lives. These resources include expected earnings and wealth accumulated through saving. But they would also include additions to wealth from rising asset prices and could include claims against the resources of others. Uncertainty about future resources is likely to encourage precautionary saving; constraints on borrowing against future resources would also tend to increase saving.

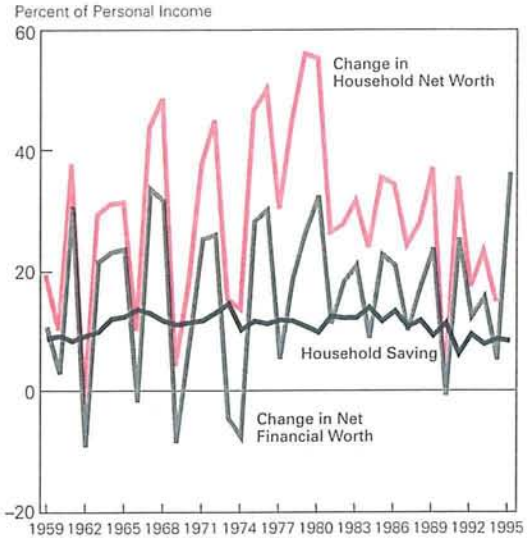
The distinction between saving and changes in wealth is especially pertinent when one looks at housing. Housing seemingly has made no contribution to saving, with additions to mortgage debt canceling out the value of new housing construction. However, the existing housing stock is very large relative to the construction taking place in any one year. Consequently, changes in the prices of existing homes can have a major effect on households' housing wealth, beyond that attributable to new construction. As can be seen in Figure 7, in most years but especially in the late 1970s, rising real estate values have added significantly to households' net wealth, even though additions to mortgage debt were roughly equal to new housing construction. (Depreciation charges against the existing housing stock, which are not shown in Figure 7, would further reduce housing's contribution to saving.)

A similar situation exists with respect to pension fund reserves, corporate equities, mutual fund shares, and some other assets. Because of capital gains—and losses—the change in the value of the asset may differ quite substantially from the increment coming from saving. Thus, over the five years from 1991 through 1995, households and nonprofit institutions disposed of almost $400 billion of corporate equities, even as the rising stock market caused the value of their holdings to increase by $2,500 billion.

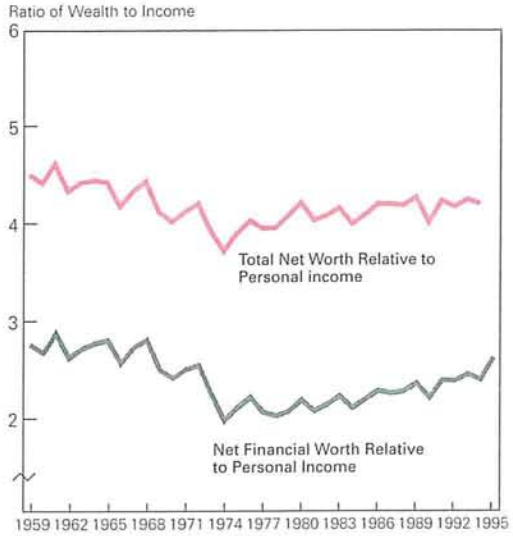
Figure 8 compares changes in financial net worth,

Figure 8A and 8B

*Household Saving and the Change in Household Net Worth*




*Household Net Worth Relative to Personal Income*




total net worth (through 1994), and saving.[13] In most years, additions to wealth surpass saving, often by a substantial margin. The growth in net worth also surpassed the growth in personal income through most of the 1980s and 1990s, so that the ratio of household net worth to income actually increased even as the saving rate fell.[14] All and all, the picture presented by net worth shows households in a stronger position than one might expect based on saving. Indeed, the life-cycle model suggests that households' strong net worth could even be the cause of the lower saving rate, with resources created through the appreciation in asset values substituting for saving.

It should also be noted that a nontrivial portion of the increase in consumption's share of income has gone to activities that are complementary to saving, specifically investment and counseling services, life insurance expenses, and various charges for services provided by financial intermediaries. Most likely, households do not consider these expenditures as consumption.

*In most years, additions to wealth surpass saving, often by a substantial margin. All in all, the picture presented by net worth shows households in a stronger position than one might expect based on saving.*

Moreover, the fact that the decline in deposits accounts for so much of the recent decline in saving raises the interesting possibility, also consistent with a life-cycle view of saving, that reductions in borrowing constraints and other increases in the efficiency of the financial system may be partly responsible for the reduction in saving. If people save in deposit accounts in order to have funds that can be quickly tapped in emergencies or because their transactions needs are erratic, then greater ability to borrow against illiquid

[13] The Flow of Funds data used in this study were obtained from two sources. Data on household saving and year-end outstanding levels of financial assets and total liabilities came from the Board of Governors of the Federal Reserve System, Flow of Funds data bases. These data were available through 1995. Data on the year-end outstanding levels of household tangible assets were taken from the June 8, 1995 Board of Governors C.9 release, *Balance Sheets for the U.S. Economy, 1945–94*. Because these data were available only through 1994, the calculation of household net worth, which relies on both financial and tangible assets as well as liabilities, could not be performed for 1995. The Board plans to release 1995 estimates of tangible assets later this year.

[14] In the 1970s, additions to wealth were larger relative to personal income, but in this high inflation period, nominal personal income grew very rapidly.

assets, such as home equity or pension reserves, and improved cash management techniques and faster availability of funds may have allowed them to economize on their deposits. Saving fell, but possibly as a by-product of a reduced need for liquidity.

And lastly, since medical problems can disrupt earnings streams and drain away assets, the existence of private health insurance plans and the knowledge that Medicare is available to help cope with the frailties of old age have removed a major source of uncertainty about the adequacy of future resources and may, therefore, have reduced the incentive to save. And for low-income individuals who are eligible for Medicaid and other forms of government assistance, the asset limitations on these programs provide a further deterrent to saving.[15]

It should not be surprising, then, that the decline in the saving rate is not as apparent to individual savers as it is to economists and policymakers. While the saving rate has fallen, the growth in households' net wealth has more or less kept pace with income. The rising stock market of the past 20 years has increased the value of corporate equity holdings and pension reserves. Rising real estate values have also contributed to households' net worth. At the same time, the availability of health insurance helps protect these assets from being depleted by medical emergencies.

While increases in asset values can substitute for saving from the standpoint of the individual, as the resulting increase in wealth can be converted into future consumption, rising asset values are a more questionable substitute for saving for society as a whole. Low saving rates are a public policy concern primarily because saving frees up resources for investment; and investment is critical to productivity growth and higher standards of living.[16] Capital gains do not make more resources available for investment and, thus, do not promise an increase in the future stream of goods and services that the economy can produce.

In particular, rising real estate values do not indicate a genuine enlargement in the stream of services that the housing stock can deliver, but simply a change in the value assigned to that stream. While individuals can tap the appreciation in their homes to increase their consumption of other goods and services, homeowners in the aggregate cannot, as a large-scale attempt to convert housing wealth into non-housing consumption would drive down home prices.

The situation for corporate equities is more ambiguous. Rising stock prices could represent an assessment that the economy's capacity to produce goods and services has increased. If so, these gains are equivalent to new investment, for an increase in the productivity of the existing capital stock, perhaps because of new technologies or new market opportunities, will increase the future stream of corporate earnings and raise the level of consumption that the economy can support. If the increase in stock prices represents an incorrect judgment and is not underpinned by an increase in productive capacity, however, there will be no increase in future consumption.

### Are Some Forms of Saving Better Than Others?

Public concern over low saving rates centers on the implications for investment. Investment provides the tools to leverage human skills and is a vehicle for introducing new technologies and organizational approaches. Thus, advocates of policies that might increase the saving rate commonly focus on the benefits of more business investment in plant and equipment, especially the latter. Housing is viewed with a jaundiced eye, largely because it is favored by the tax system but also because it is not seen as enhancing the productive capability of the economy. And no advocates of more saving are calling for more spending on consumer durables. Yet saving by way of investing in consumer durables may also push out our production frontier, or at least did so in the past. The widespread acquisition of labor-saving household appliances played an important role in enabling women to enter the workplace, and the mobility provided by motor vehicles has vastly broadened the range of employment options that workers can consider. By permitting a fuller utilization of human capital, investment in these consumer durables has increased our productive capacity.

[15] See Hubbard, Skinner, and Zeldes (1994).

[16] Countries with low saving rates must either forgo these investment opportunities or depend upon attracting the savings of other countries, which then enjoy the return to these investments. The host country will still benefit from the employment generated by the investment activity and from any externalities the investments throw off. These externalities might take the form of the introduction of new technologies or management techniques. The source of the saving, however, commands the income stream generated by the investment.

Individuals will still be able to convert capital gains into future consumption by selling their assets, but society as a whole cannot.

## *Conclusion*

Over the past 10 years, the personal saving rate has fallen sharply. This decline occurred in the context of federal government dissaving, which has been driven largely by rising transfer payments, boosting personal income relative to GDP. Out of this larger income, a smaller fraction has been saved.

Although some economists have argued that government dissaving will be offset by increased private saving, this outcome is less puzzling when one considers that the growth in transfer payments during this period was propelled, in large part, by escalating medical expenditures. Through Medicare and Medicaid, the government ensures that certain segments of the population, the elderly and the poor, can achieve a level of medical consumption beyond what their individual incomes would permit. At the same time, employer contributions for private health insurance and the expenditures financed by this insurance have grown dramatically.

Thus, the bottom line is that the answer to the question—where did the saving go?—is medical care. Rising expenditures on medical services are absorbing a growing fraction of income. Thus, the saving problem is not about thrift versus profligacy, good versus bad; rather, it is a competition between two "goods"—more and better medical care, on the one hand, and more investment, on the other.

Too often, discussions of the saving rate present saving itself as a "good" and, thus, any decline in saving as undesirable. But saving is a means to an end. For the individual, saving is a means of insuring against the unexpected mishap and smoothing consumption over time. But it may be possible to achieve these objectives in other ways—through insurance programs or increasing the efficiency of financial markets or tapping capital gains.

What is eminently rational behavior for the individual, however, may impose costs on society as a whole, if the curtailment in saving translates into a reduction in investment that affects the economy's future capacity to produce and consume. Trying to stimulate personal saving through changes in the tax system or other incentives is one policy response. Reducing the federal deficit could also make more resources available for investment; and shifting the composition of investment to more productive forms could increase the societal return associated with a given volume of saving. But the goal should be higher standards of living, with more saving one of several means to that end.

## *References*


Attanasio, Orazio P. 1993. "A Cohort Analysis of Saving Behavior by U.S. Households." National Bureau of Economic Research Working Paper No. 4454.

Bosworth, Barry, Gary Burtless, and John Sabelhaus. 1991. "The Decline in Saving: Evidence from Household Surveys." *Brookings Papers on Economic Activity*, no. 1, pp. 183–253.

Gokhale, Jagadeesh, Laurence J. Kotlikoff, and John Sabelhaus. 1995. "Understanding the Postwar Decline in United States Saving: A Cohort Analysis." Federal Reserve Bank of Cleveland Working Paper 9518.

———. 1996. "Understanding the Postwar Decline in United States Saving: A Cohort Analysis." National Bureau of Economic Research Working Paper No. 5571.

Hubbard, Glenn, Jonathan Skinner, and Stephen P. Zeldes. 1994. "Expanding the Life-Cycle Model: Precautionary Saving and Public Policy." *The American Economic Review, Papers and Proceedings*, vol. 84, no. 2, May, pp. 174–79.

Kopcke, Richard W., Alicia H. Munnell, and Leah M. Cook. 1991. "The Influence of Housing and Durables on Personal Saving." *New England Economic Review*, November/December, pp. 3–16.

Munnell, Alicia H. and Leah M. Cook. 1991. "Explaining the Postwar Pattern of Personal Saving." *New England Economic Review*, November/December, pp. 17–28.

Summers, Lawrence and Chris Carroll. 1987. "Why Is National Saving So Low?" *Brookings Papers on Economic Activity*, no. 2, pp. 607–42.

U.S. Bureau of Economic Analysis. 1992. *National Income and Product Accounts of the United States, volume 2, 1959–88*. Washington, D.C.: Government Printing Office.

———. 1996. "Improved Estimates of the National Income and Product Accounts for 1959–95: Results of the Comprehensive Revision." *Survey of Current Business*, January/February, pp. 1–31.

# The Boston Fed Has a Home on the World Wide Web

The Federal Reserve Bank of Boston has established a site on the World Wide Web (WWW). The new "home page" is a gateway to Federal Reserve economic and statistical information and offers on-line access to most Boston Fed publications.

Guests to the site can learn more about the Federal Reserve System and its operations in Boston. There is information specifically for users of the Bank's financial services and information of interest to the general public. The site also provides many useful links to other sites.

The following Research publications are featured on the new site:

- New England Economic Review
- Regional Review
- New England Banking Trends
- Fiscal Facts
- New England Economic Indicators

**World Wide Web address: http://www.bos.frb.org**

# Do Municipal Bond Yields Forecast Tax Policy?

*Peter Fortune*

*Senior Economist and Advisor to the Director of Research, Federal Reserve Bank of Boston.*

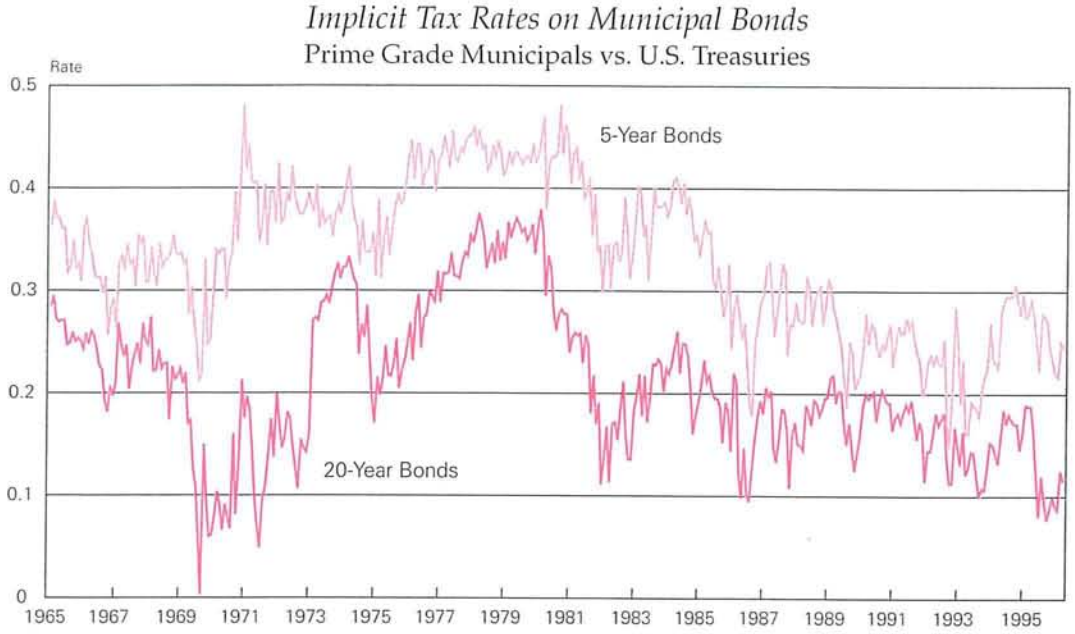
During the 1995–96 Presidential primary season, the "flat tax" was a widely discussed tax reform. In July of 1995, a flat tax plan was proposed by House Majority Leader Dick Armey and Senator Richard Shelby (H.R. 2060). The Armey-Shelby flat tax was subsequently adopted as the key component of the Forbes campaign and, with modifications, as a board in the platforms of other Presidential contenders.

The Armey-Shelby flat tax plan proposed a single tax rate on all taxable income and a redefinition of taxable income to exclude interest and dividends. While this would have important effects on the valuation of many financial instruments, the impact on the market for municipal bonds could be profound. The exclusion of interest income from the tax base would eliminate the tax advantage that municipal bonds traditionally have enjoyed. The effect on relative interest rates is clear: The end of federal taxation of interest income should create a new security market equilibrium in which municipal bond rates would equal the rates paid on equivalent taxable securities. This entails a sharp rise in the *relative* yield on municipal bonds. However, some proponents of the flat tax argue that the general level of interest rates would fall, so that the cost of capital for municipalities might fall even though the relative yield on municipal bonds rises.

During the time of the flat tax debate, the interest rate on newly issued municipal bonds rose relative to the rate on U.S. Treasury bonds. This was frequently cited as arising from expectations that a flat tax might be implemented. For example, in the May 13, 1996 issue of its *Fixed Income Research Relative Value Report*, Lehman Brothers reported that "After spending most of 1995 in fear of a flat tax, the municipal bond market has rallied relative to taxables for the first several months of 1996."

That security prices and yields reflect anticipation of relevant future events is an axiom of financial economics. In its extreme form—the Efficient Markets Hypothesis—current asset prices reflect an optimal

Figure 1




Source: Salomon Brothers and author's calculations.

forecast of the future cash flows generated by the asset. Even if one is not an efficient markets enthusiast, the partial reflection of anticipations in asset prices is widely accepted. Most of the work in this area has been done in the context of stock and commodity markets and, as we shall see, very little has been done for the municipal bond market.

Throughout this article, we will use the concept of the "implicit tax rate," also called the "breakeven tax rate" because an investor with that tax rate will be indifferent between a tax-exempt bond and a taxable bond, and the "implicit subsidy rate" because it measures the subsidy to municipal borrowing arising from tax exemption. If $R_m$ and $R$ denote the yields on municipal bonds and on taxable bonds with the same term to maturity and credit risk, and if $s$ denotes the implicit tax rate, then the implicit tax rate is calculated as $s = 1 - R_m/R$.

Figure 1 shows the implicit tax rate reflected in interest rates for 5-year and for 20-year bonds. The 5-year implicit tax rate is calculated using the definition given above and Salomon Brothers 5-year Prime Grade general obligations (GOs) and 5-year U.S. Treasury bond yields; the 20-year implicit tax rate uses the Salomon Brothers data for 20-year bonds. A clear correlation can be seen between the implicit tax rates and the personal income tax rates paid by affluent investors. For example, for 5-year bonds the implicit tax rate fell after the Economic Recovery and Tax Act of 1981 (ERTA), which reduced tax rates paid by upper-income investors. It fell again after the Tax Reform Act of 1986, which further reduced the tax rates paid by the affluent and also eliminated many competing tax shelters. The same pattern is shown for 20-year bonds. Note also that the implicit tax rate fell to low levels in 1969, particularly for 20-year bonds. This was attributed at the time to a debate about including municipal interest income in the new Alternative Minimum Tax.

If the implicit tax rate has no information content, implicit tax rates at any particular time should be uncorrelated with the path of actual tax rates as they evolve after that time. At the other extreme, if the implicit tax rate has its maximum information content, it should perfectly predict future tax rates, with due allowance for statistical noise. What is the information content of implicit tax rates as predictors of future tax rates? The answer to that question will let us know how much credence to attach to tax rate debates when assessing changes in relative interest rates. It can also

tell us something about the efficiency of security markets. The goal of this study is to assess the evidence bearing on this question.

The article is structured as follows. The first section briefly reviews other studies of the connection between implicit and actual tax rates. The second section addresses the question of how anticipations of future tax rates should be measured. The third section develops and estimates a model of the information content of implicit tax rates. This model allows the information content to change over time as tax policy debates wax and wane. The article ends with a brief summary.

## *I. Previous Research on Tax Policy and Implicit Tax Rates*

Our study will examine the connection between the expected future tax rate of a representative investor and the implicit tax rate. A key question in this analysis is who should represent the representative investor. That is, whose tax rates matter? Following a brief discussion of this question, we will look at some empirical evidence on the relationship between tax policy events and bond yields.

### *Whose Tax Rates Matter?*

It will come as some surprise to noneconomists that there has been a debate about whether implicit tax rates are affected at all by personal income tax rates. In 1977, Eugene Fama and Merton Miller independently developed models in which the implicit tax rate is determined by the corporate income tax rate. In this "new view," personal income taxes do not matter at all in the determination of municipal bond yields. If this is true, the representative investor should be a corporation facing the statutory corporate income tax rate.

Fama's explanation of the new view rests on the assumption that corporations, particularly commercial banks, are the marginal investors in municipal bonds. Because they can borrow at the corporate bond rate and deduct their interest payments, the after-tax cost of funds is $(1 - \tau_c)R$, where $\tau_c$ is the corporate tax rate and $R$ is the taxable bond rate. If the opportunity to invest in municipal bonds is unrestricted they will hold municipal bonds when $R_m$ exceeds $(1 - \tau_c)R$ and taxable bonds when $R_m$ is less than $(1 - \tau_c)R$. Security market arbitrage will ensure that $R_m = (1 - \tau_c)R$ in equilibrium as banks borrow at the taxable rate and invest at the tax-exempt rate. Thus, the banking sector's demand for tax-exempts is infinitely elastic at the after-tax interest rate on taxable bonds, and banks are the marginal investors in municipal bonds. The implicit tax rate will be $\tau_c$.

Miller's exposition of the new view sees nonfinancial corporations as the marginal suppliers of municipal debt. This might be motivated by corporate access to the municipal bond market for industrial development, real estate development, environmental, and other purposes. But Miller's story does not require direct corporate access to municipal financing. It rests on the assumption that municipal bonds and common stock are very close substitutes. Suppose that they are perfect substitutes. Investors in corporate equity (shareholders) will require a marginal pre-tax return on capital of $R_m/[(1 - \tau_c)(1 - \tau_e)]$, where $\tau_e$ is the personal income tax rate for equity income (dividends or capital gains). This is the municipal bond rate grossed up by the taxes paid at both corporate and personal levels, and it is the pre-tax return on plant and equipment required to give shareholders an after-tax return equal to the municipal bond rate. Investors in corporate debt (bondholders) will require a return on corporate capital equal to the taxable bond rate, $R$, so that is the marginal pre-tax cost of capital required for debt finance. The optimizing capital structure is that for which the marginal costs of capital for debt and equity are equal, hence an equilibrium exists when $R_m = [(1 - \tau_c)(1 - \tau_e)]R$. Miller also assumes that the personal income tax rate on equity income is zero, so the equilibrium yield relationship is $R_m = (1 - \tau_c)R$.[1] The implicit tax rate will, as in Fama's model, be the corporate income tax rate but the reason is that corporations are the marginal suppliers of a perfect substitute for municipal debt.

Whatever the merits of the new view when it was first presented in the 1970s, subsequent events have weakened, if not eliminated, its validity. During the 1980s, the deductibility of interest paid to carry tax-exempt bonds was eliminated. Banks lost their incentives to invest in municipal bonds, and dominance of the municipal bond market shifted from banks to individual investors, primarily through mutual funds. This ended the foundation on which Fama's version of the new view rested. At the same time, changes in

[1] A zero personal income tax rate is not farfetched. Dividend-paying stocks tend to be held by tax-exempt financial institutions, and low-dividend stocks are affected by the capital gains tax rate. The effective capital gains tax rate can be quite low because realization can be deferred as long as the shareholder wishes.

Table 1
*Events Expected to Affect Implicit Tax Rates*

| Predicted Sign | Regression Results: Estimated Impact on Implicit Tax Rate 5-yr | 10-yr | 20-yr | Date | Event |
|---|---|---|---|---|---|
| + | c | c | c | Jan67 | President Johnson proposes tax surcharge |
| + | c | c | x | May–Jun68 | One-year surtax passes and is signed |
| – | c | c | c | Mar69 | Several members of House Ways and Means propose changing taxation of municipal interest |
| – | c,s | c,s | c,s | Jul69 | House Ways and Means passes minimum tax proposal including municipal interest in tax base |
| + | c | c | x | Sep69 | Senate Finance Committee members oppose including tax-exempt income in minimum tax |
| + | x | x | x | Dec69 | Surtax extended to 1970 |
| + | c,s | c | c | Jan70 | President Nixon proposes revenue-sharing plan |
| + | c,s | c,s | c,s | Nov–Dec70 | Proposed revenue-sharing plan expanded |
| + | c | c | x | Mar72 | House approves revenue-sharing |
| – | c | c,s | c,s | July74 | House Ways and Means passes bill reducing top personal income tax rate to 50% |
| – | x | c | c,s | Nov–Dec74 | New York City financial crisis |
| + | c | x | c | Jan78 | President Carter proposes upper-income tax rate increase |
| – | c | c | c | Jun78 | Proposition 13 passes in California |
| – | x | x | c | Jun80 | Candidate Reagan proposes large cuts in income tax rates |
| – | x | x | c | Nov80 | Reagan elected President |
| – | c | c | c | Dec80 | IRS rules commercial banks cannot deduct interest paid for carrying municipal securities |
| – | c | c | c | Jan81 | President Reagan proposes tax cut |
| – | c | c | c | Aug81 | Economic Recovery Tax Act passes |
| – | x | x | x | Aug82 | Tax Equity and Fiscal Responsibility Act allows deduction of only 85% of interest paid to carry municipal bonds |
| – | c | c | c,s | Apr83 | Social Security Act amendment exposes municipal interest to taxation |
| – | x | x | x | Mar–May83 | Washington Public Power Supply default |
| – | c | c | c | Aug84 | Deficit Reduction Act reduces bank interest deduction for carrying municipals to 80% |

the tax code eliminated the foundations for Miller's version of the new view. For example, after 1986 (and until 1993) the maximum personal income tax rate was less than the corporate income tax rate, so personal income taxation would become relevant.

The early empirical evidence for the new view appeared to be favorable. In a widely cited paper, Trczinka (1982) found that the implicit tax rate was very close to the corporate income tax rate during the 1970s. However, later evidence has led to a rejection of the Fama-Miller view. Fortune (1988) showed that Trczinka's results were unique to the 1970s, and that during the 1980s the movements in the implicit tax rate were consistent with changes in personal income tax rates. This shift is, of course, consistent with the tax code changes during the 1980s, which weakened the new view's foundations. In addition, as Figure 1 shows, the implicit tax rate fell sharply in 1981, when ERTA reduced marginal tax rates of the affluent but did not change the corporate tax rate. This event occurred before the tax code changes that eliminated the new view's validity, suggesting evidence against that view even in its prime.

Thus, we adopt the position that the representative investor is an individual for whom the personal income tax is the tax-related force driving his portfolio decisions. Our results will support that choice.

### *Event Studies*

Event studies also indicated the importance of personal income taxes. These event studies were the first to show that the implicit tax rate did reflect information about future tax rates. Poterba (1986,

Table 1 continued

*Events Expected to Affect Implicit Tax Rates*

| Predicted Sign | Regression Results: Estimated Impact on Implicit Tax Rate 5-yr | 10-yr | 20-yr | Date | Event |
|---|---|---|---|---|---|
| – | c | c | c,s | Nov84 | Reagan reelected, Treasury I tax proposal announced |
| – | c | c | c | Dec85 | House Ways and Means passes bill incorporating major tax reforms including reducing tax rates and tax shelters |
| – | x | x | c | Mar86 | Senate Finance Committee considers including municipal interest in alternative minimum tax |
| – | c,s | c,s | c,s | May86 | Senate Finance Committee passes revised version of the Dec '85 House bill |
| – | x | x | x | Sep86 | Tax Reform Act passed |
| – | x | c | x | Apr88 | In *South Carolina vs. Baker*, Supreme Court rejects constitutional foundation for tax exemption |
| – | x | x | c | Nov90 | President Bush signs Revenue Reconciliation Act, eliminates "bubble" tax rate, sets cap on capital gains tax rate |
| + | c | x | x | Aug93 | President Clinton's Omnibus Budget Reconciliation Act passed, increases tax rates for higher-income groups, treats market discounts as ordinary income |
| – | c | c | x | July–Dec95 | Armey-Shelby Flat Tax (HR 2060) Introduced and debated |
| – | c | x | x | Jan–Feb96 | Several Presidential primary candidates endorse flat tax |
| Summary | 5-yr | 10-yr | 20-yr | Total | |
| Number c | 22 | 21 | 21 | 64 (66.7%) | |
| Number c and s | 4 | 4 | 7 | 15 (15.6%) | |
| Number x | 10 | 11 | 11 | 32 (33.3%) | |
| Number x and s | 0 | 0 | 0 | 0 | |

The regression equation was an MA(1) with the change in the implicit tax rate as the dependent variable.
c denotes regression coefficient has the predicted sign.
s denotes regression coefficient is statistically significant (5%, one-tail).
x denotes regression coefficient has incorrect sign.

1989) investigated the movements in implicit tax rates during important tax policy events. In his first paper, Poterba included 22 tax policy events and seven credit risk events in the 1962–82 period. In his second paper, he focused on 13 tax policy events between 1968 and 1988. Rather than report his results, we have used the same methods on an updated list of 32 events in the 1967–96 period.

An updated list of relevant events is shown in Table 1. A multiple regression was estimated in which the implicit tax rate was regressed on dummy variables for specific tax policy events.[2] This was done for implicit tax rates on 5-year, 10-year and 20-year bonds. The interest rates used were the Salomon Brothers monthly series for Prime Grade GOs and U.S. Treasury bonds. The first column of the table reports the predicted sign of the coefficient on the event. The next three columns summarize the regression results (c indicates a correct sign, x indicates an incorrect sign and s indicates statistical significance). The remaining two columns show the date and a brief description of the event. There are 96 event coefficients (32 tax policy events times 3 bond maturities).

Of the 96 event effects estimated, two-thirds of the coefficients have the correct sign, giving the appearance that tax policy events affect the implicit tax rate. However, only 16 percent of the coefficients have both the correct sign and statistical significance, weakening

[2] A dummy variable is a variable designed to reflect the existence of a condition. It is defined as one during the period the condition exists, and zero otherwise. Each tax policy event was reflected by a dummy variable having a value of one during the month(s) of that event, and zero otherwise.

the conclusion that tax policy events are embedded in implicit tax rates. Only four tax policy events showed statistically significant *and* correct signs for two or more of the regressions. These events were the July 1969 passage of a House Ways and Means Committee bill including municipal interest in the proposed minimum tax, the late 1970 expansion of the revenue-sharing plan proposed by President Nixon, the July 1974 House Ways and Means passage of a bill reducing the top personal income tax rate to 50 percent, and the May 1986 Senate Finance Committee passage of the House version of the Tax Reform Act.

*Event studies were the first to show that the implicit tax rate did reflect information about future tax rates, but an updated event study provides little comfort for those who believe that implicit tax rates carry abundant information about future tax policy.*

During much of the period the municipal bond market was segmented, with commercial banks dominating the market for shorter maturities and individuals dominating the longer-term market. This segmentation might explain why the results for the 5-year implicit tax rate differed from the 10- and 20-year results. Thus, while the July 1974 House Ways and Means bill reducing top personal income tax rates did depress implicit tax rates at all maturities, the effect was statistically significant only for the 10- and 20-year terms, where high-income individuals were the primary investors. The 1983 Social Security Act Amendments, which exposed municipal interest received by high-income Social Security beneficiaries to taxation, affected all three implicit tax rates in the correct direction, but only the 20-year effect was significant. However, other events that should have affected the short end where banks invested did not have the expected effects. For example, the 1982 Tax Equity and Fiscal Responsibility Act (TEFRA) and the 1984 Deficit Reduction Act both reduced bank incentives to hold municipal bonds. TEFRA showed incorrect signs in all three cases, while the milder DRA showed the correct sign in all three cases.

Some events that should be important for implicit tax rates are not. For example, President Reagan's January 1981 tax cut proposal and passage of the Economic Recovery Tax Act in August of 1981 show correct signs, but neither of these effects is statistically significant. Both the South Carolina vs. Baker decision in April 1988, and President Clinton's signing of the 1993 Omnibus Budget Reconciliation Act in August of 1993, show incorrect signs in two of the three cases. Of particular interest, the introduction and debate of the Armey-Shelby flat tax bill showed correct signs in two of the three regressions, but neither was statistically significant, and the 1996 Presidential primary debates over the flat tax showed correct signs in only one of the three cases, again with no statistical significance.[3] Thus, Lehman Brothers' attribution of high municipal bond yields, and low implicit tax rates, to flat tax debates is not supported.

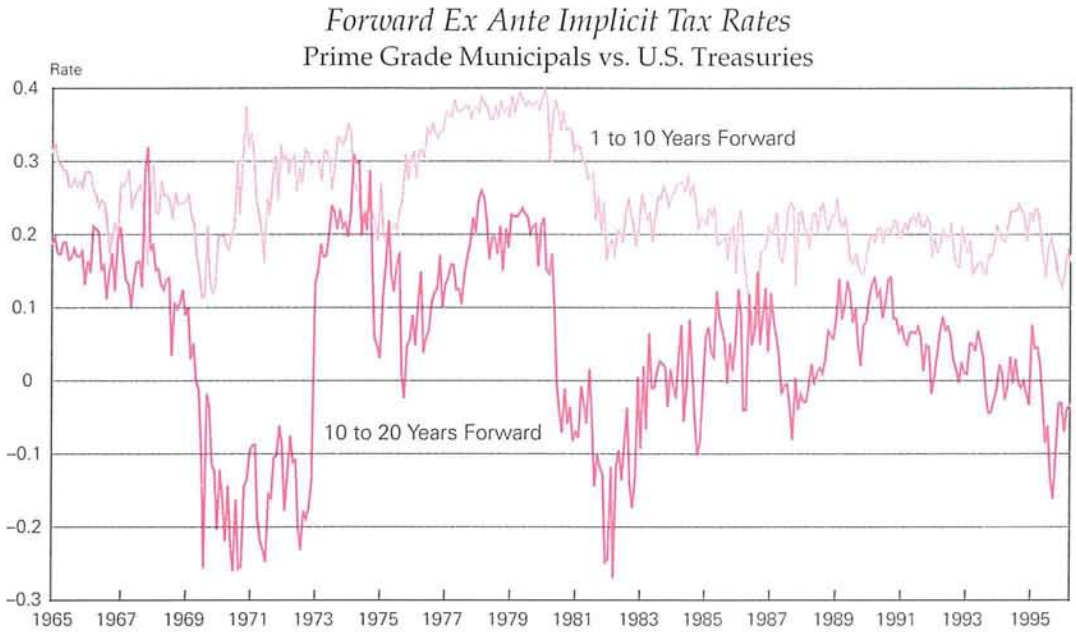
The general failure of this event study to support the view that implicit tax rates contain a great deal of information about tax policy is somewhat unsettling, though event studies do have a number of well-known problems. The selection of an "important event" is in the eye of the beholder, and the timing of economic events is notoriously hard to identify because the real question is not when did legislation pass, or when was it proposed, but when did expectations about future tax rates change. Furthermore, the best event study can only tell whether there was a correlation between the timing of an event and related market activity or prices; it says nothing about the magnitude of the event's impact, or about the dynamics surrounding that impact. Even so, our update of Poterba's event study provides little comfort for those who believe that implicit tax rates carry abundant information about future tax policy.

## *II. Measuring Expected Future Tax Rates*

In order to examine the connection between implicit tax rates and expectations of future tax rates, we must have some measure of expectations. In this section we consider two approaches to deriving measures of expected future tax rates. We first assess the

[3] We date the flat tax debate in two events: the period in 1995 following introduction of H.R. 2060 (July–December) and the period of the Presidential primary debate over the flat tax (January–February 1996).

Figure 2




Source: Salomon Brothers and author's calculations.

tax rates implicit in the yield curves for taxable and tax-exempt bonds. The result, called the *ex ante implicit tax rate*, has been used by Fortune (1991) in an analysis of the determinants of municipal bond yields and, more recently, by Park (1995) in an examination of the relationship between federal financial conditions and expected tax rates. Rejecting this measure of expected future tax rates, we develop an alternative approach by determining the future tax rates that investors actually pay, using the discounted value of these as a measure of expected future tax rates. The result, called the *ex post tax rate*, measures the value of taxes actually avoided over the life of a municipal bond.

### *The Ex Ante Implicit Tax Rate*

Among the earliest studies of ex ante implicit tax rates is Kochin and Parks (1988), which used the yield curves for municipal and U.S. Treasury bonds to calculate the ex ante tax rates implicit in the yield curves for taxable and tax-exempt bonds. If $s_{t,t+k}$ is the spot implicit tax rate at time $t$ on bonds maturing at time $t + k$, $s_{t,t+m}$ is the spot implicit tax rate at time $t$ on bonds maturing at a more distant time, $t + m$, and $T_{k,m}$ is the ex ante implicit tax rate at time $t$ for the interval $t + k$ to $t + m$, then, Kochin and Parks show, the ex ante implicit tax rate is a weighted average of the two spot implicit tax rates. Specifically,

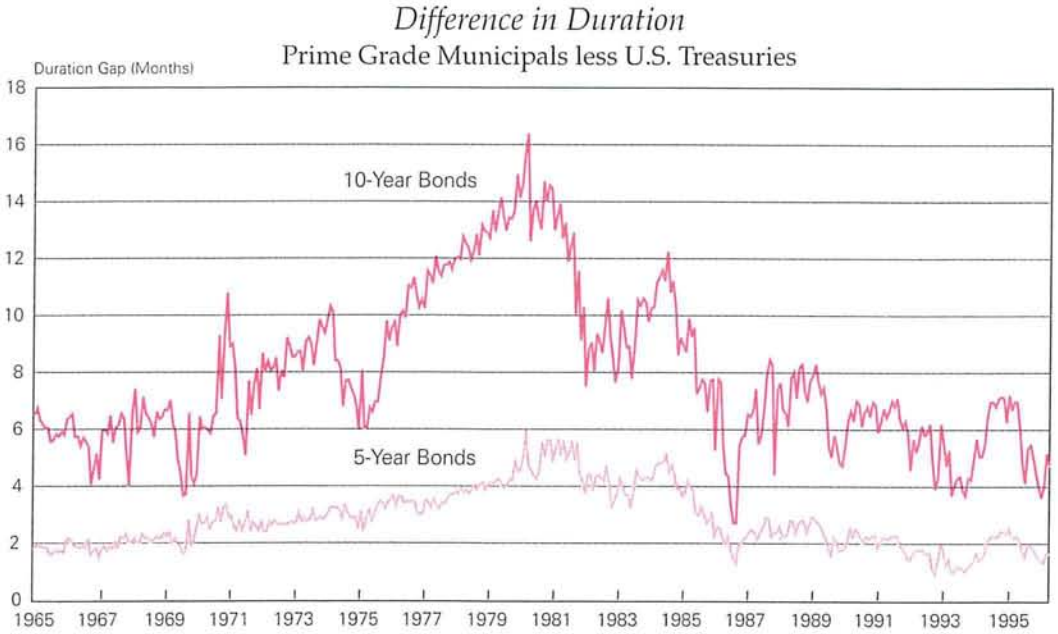
$$T_{k,m} = w_t s_{t,t+m} + (1 - w_t) s_{t,t+k},$$

$$\text{where} \qquad w_t = \sum_{i=1}^{m} (1 + R_i)^{-i} / \sum_{i=k+1}^{m} (1 + R_i)^{-i}. \qquad (1)$$

Figure 2 shows the ex ante implicit tax rates for two future intervals, 1 to 10 years into the future, and 10 to 20 years into the future.[4] The 1- to 10-year ex ante implicit tax rate appears to conform roughly with expectations: It is always positive and it ranges between 0.2 and 0.4. It also shows a tendency to peak just before major tax rate reductions (the 1981 ERTA and the 1986 TRA). However, the 10- to 20-year ex ante implicit tax rate makes no sense. It is negative for extended periods, and it seems too low in the 1980s.

[4] The Kochin-Parks calculation of the weight, $w_t$, requires information on the zero-coupon yield curve, showing the interest rate on single payments through the life of the longest bond. Because these data are not available, we use the yield curve for coupon-bearing municipal bonds.

Figure 3




Source: Salomon Brothers and author's calculations.

This pattern is the same as that found by Kochin and Parks.

The negative ex ante implicit tax rates in the distant future present a serious problem for the interpretation of relative yields as driven by tax rate expectations. Clearly, investors in 20-year municipal bonds do not expect negative future tax rates 10 to 20 years out. Furthermore, the observation of negative ex ante implicit tax rates should elicit arbitrage behavior that eliminates the phenomenon. For example, investors with high tax rates can earn riskless profits by buying long-term municipal bonds, receiving the high municipal bond yields, and selling Treasury bonds with an equal duration. This would reduce municipal bond yields and raise Treasury bond yields, restoring the positive ex ante implicit tax rate. However, Figure 2 shows that negative ex ante implicit tax rates last for extended periods, suggesting that this arbitrage does not function, or that there is some other explanation for the negative implicit tax rates.

A more plausible interpretation of Figure 2 is that the bond yields from which the ex ante tax rates are constructed are contaminated by factors other than tax-rate expectations, and that this contamination becomes more serious as we look further into the future. These "contaminants" all can be capsulized in a generic "risk premium" that investors attach to municipal bonds. This risk premium is the extra return required to compensate for all nontax differences between municipal and Treasury bonds.

Several candidates come to mind. The first is duration. Municipal bonds and U.S. Treasury bonds with the same term to maturity necessarily differ in their durations and, therefore, in the interest rate risk their holders experience. The reason is that the coupons on newly issued municipal bonds are less than the coupons on Treasury bonds with the same maturity. This shifts the distribution of cash payments further into the future for municipal bonds than for Treasury bonds, making the duration longer for municipals. Figure 3 shows, for 5-year and 10-year terms, the difference between the durations of Prime Grade GOs and U.S. Treasury bonds. The duration of municipal bonds is always greater than the duration of Treasury bonds because tax exemption shifts the stream of cash payments further into the future: Both bonds receive the face value (say, $1000) at maturity, but municipals pay a lower coupon. Figure 3 also shows that the duration gap widens when interest rates are high and narrows when they are low: When

interest rates rise, the duration of Treasury bonds falls relative to the duration of municipal bonds. Thus, we would expect that the ex ante implicit tax rates are a biased estimate of future tax rates because of duration differences in the instruments, and that this bias increases with interest rates. If a positive relationship exists between duration and yields required by investors, this bias will induce higher municipal bond yields for more distant maturities and contribute to the very low, sometimes negative, implicit tax rates on 20-year municipal bonds shown in Figure 2.

A second contaminant is callability. U.S. Treasury bonds do not carry call features, but many municipal bonds allow the issuer to call the bond at its discretion after some period of time. Because the call option held by the issuer has value, the municipal bond buyer, who is the option's writer, pays a lower price for a callable bond than for an equivalent non-callable bond. This shows up as a higher interest rate on the municipal bond and, therefore, as a lower implicit tax rate. Again, the effect is to make the implicit tax rate a downward-biased measure of expected future tax rates. The effect of callability is largely confined to municipal bonds with over 10 years to maturity. Cook (1982) reports that in a sample of 860 municipal bonds issued in 1977–78, less than 1 percent of 5-year bonds were callable, and only about 10 percent of 10-year bonds could be called before maturity. In contrast, 88 percent of 15-year bonds and 98 percent of 20-year bonds had call features.

*The "risk premium" that investors attach to municipal bonds is the extra return required to compensate investors for such nontax differences between municipal and Treasury bonds as duration, callability, and credit risk.*

The value of the call feature is not constant, so the bias in implicit tax rates is not constant. Two important determinants of the call's value are the volatility of interest rates and the time to first call. Because the call's value increases with volatility, the bias will be greater at times of high interest rate volatility. Because the call's value decreases as the time to first call gets closer, the bias will be smaller for bonds with a short time to first call. It is, perhaps, not coincidental that the negative implicit tax rates for 20-year bonds shown in Figure 2 were in periods of considerable financial market volatility, the early 1970s and the early 1980s.

The call option problem can be addressed in several ways. One is to explicitly measure the value of the call option using one of the many models available for valuing options on bonds. This would lead one to calculate option-adjusted spreads, and to use these interest rates in computing both spot and ex ante implicit tax rates. A second and simpler approach, which we follow, is to eliminate from our analysis bonds with more than 10 years to maturity, on the grounds that these are the ones most tainted by callability.

Yet a third contaminant of ex ante tax rates is credit risk. The probability of a U.S. Treasury default is negligibly low, although recent fiddling with debt limits might suggest some caution. Municipal bonds are another matter. Most tax-exempt bond defaults have been confined to revenue bonds, such as the Washington Public Power Supply default in 1983. Very few general obligation bonds have defaulted, but the possibility always remains that even Prime Grade GOs can default. Because this is more likely to occur in the distant future, default risk prospects will be most prominent for terms beyond five years. Indeed, default risk is one of the reasons often given for the decline in the implicit tax rate as bond maturity gets more distant (Figure 1).

A recent paper by Chalmers (1995) suggests that default risk does not explain the inverse relationship between the implicit tax rate and time to maturity revealed in Figure 1. Chalmers examines advance refunding (defeasance) of municipal bonds. An advance refunding involves determining the portfolio of U.S. Treasury securities required to match the cash payments to be made on outstanding municipal bonds, then issuing a new municipal bond in an amount equal to the cost of this portfolio. Because the proceeds are invested in an irrevocable escrow of the requisite U.S. Treasury bonds, from which the income and principal are used to make the payments on the refunded issue, the defeased municipal bond has no default risk. The incentive to do this is, of course, the reduction in debt service achieved by taking advantage of the higher yield on Treasury securities and substituting a new, lower-coupon municipal bond for the old bond.

Chalmers found that the inverse relationship between maturity and implicit tax rate held for defeased bonds as well as for the bonds used in the Salomon Brothers Prime Grade GO series. Furthermore, the rate of decline was about the same for both series. Because defeased bonds have no default risk and no call risk (the call option is typically extinguished at defeasance), Chalmers' results suggest that the relationship between implicit tax rates and maturity derives from factors other than default risk and callability. One possibility is that it reflects the uncertainty about future tax rates.

*We reject ex ante implicit tax rates as measures of expected future taxes because they appear to be a poor measure.*

We reject ex ante implicit tax rates as measures of expected future taxes because they appear to be a poor measure. We have seen that the ex ante implicit tax rates constructed from the Salomon Brothers bond yield data are not determined by tax rate anticipations alone, but are also affected by duration, callability, default risk, and market risk differentials (though Chalmers' work suggests these effects might be less severe than previously thought).

### *Ex Post Tax Rates*

An alternative measure of expected future tax rates can be constructed from actual future tax rates. This entails a minor truncation: Ex post tax rates cannot be calculated for periods in which actual future tax rates are not known. For example, because we have data up through 1995, our series for the ex post tax rate on 5-year bonds cannot be extended beyond 1990; the tax rate on 10-year bonds cannot be calculated after the end of 1985.

Suppose that we have data on a representative investor's actual tax rates after time $t$. Let $\tau_{t+i}$ represent the investor's tax rate in period $t + i$. Noting that coupons are typically paid semiannually, we can construct the time series for tax rates paid at each future semiannual period. This series is denoted as $\tau_{t+6}, \tau_{t+12}, \tau_{t+18}, \ldots, \tau_{t+6n}$, with $t$ being the spot date and $n$ being the number of semiannual periods to maturity. The average ex post tax rate ($T_t$) can then be constructed from known statutory tax rates over the period to maturity. We use the following definition, in which the actual semiannual ex post tax rates are discounted to the present using the weights derived in Box 1.

$$T_t = \sum_{i=1}^{N/12} w_i \tau_{t+6i},$$

$$\text{where} \quad w_i = R_m(1 + R_m)^{-i}/[1 - (1 + R_m)^{-N}]. \quad (2)$$

No single ex post tax rate applies to all time horizons and all investors. The ex post tax rate will be different for each bond maturity because the horizon over which future tax rates are considered changes. It will also vary across investors, because an investor's income path determines his tax rates. Finally, it will vary over time, as statutory tax rate schedules change.

We have constructed the ex post tax rate for several representative investors over our sample period (1965–95). Because our focus is on the implicit tax rates derived from municipal bonds, we have focused on high-income individuals who will invest in these securities. Our analysis is based on federal income tax rates, excluding state income taxes which might be relevant if a state's municipal bonds were sold entirely within the issuing state.[5]

Figure 4 shows the actual tax rates paid by three representative investors: those who pay the maximum federal personal income tax rate, and those who earn real taxable incomes (1980 dollars) of $100,000 and $50,000. The figure shows that prior to the 1986 Tax Reform Act a pronounced difference existed in marginal tax rates, reflecting the progressivity of the tax code. The maximum tax rate was 70 percent prior to 1982, about 10 to 20 percentage points more than the marginal tax rate associated with $100,000 of real taxable income. The Economic Recovery Tax Act of 1981 greatly reduced the differentials, and the 1986 Tax Reform Act reduced them even more. A reemergence of tax rate differentials began in 1993, with

[5] Most states include tax-exempt interest in the state's definition of taxable income only if the interest is paid by an out-of-state issuer. This gives an extra tax advantage to in-state investors. However, whether this affects the yield on a state's municipal bonds depends on whether out-of-state investors buy those bonds. If they do, the state income tax plays no role in the pricing of the municipal bonds because the marginal investors are out-of-state. Only when the marginal investors are in-state will the state's income tax rate affect the yield on municipal bonds.

**Box 1: The Implicit Tax Rate and Ex Ante Implicit Tax Rates in an Efficient Market**

Consider a newly issued N-period taxable bond with a coupon rate equal to the required interest rate, therefore priced at par. Let $R$ be the yield to maturity (also the coupon rate) on the taxable bond, and $R_m$ be the yield to maturity (and coupon rate) on an N-period tax-exempt bond. Also, let $\tau^*_{t+i}$ be the tax rate presently anticipated to apply in period $t + i$, and $\theta_{t+i}$ be the marginal risk premium required on income from a taxable bond in period $t + i$, required to make that income equivalent to income from tax-exempt bonds with the same term. The expected tax rate is conditional on the information available at the present time, denoted as $\Omega_t$. Thus, $\tau^*_{t+i} = E(\tau_{t+i} | \Omega_t)$.

The sequence of marginal risk premia $\{\theta_1, \theta_2, \ldots, \theta_N\}$ is, of course, unobservable. The "risk premia" arise for any reason that causes investors to attach different values to an expected stream of after-tax income from taxable bonds and an equal expected stream of income from tax-exempt bonds. For example, differences in call features, in duration, in underlying collateral, and in the general quality of the bonds can be captured in the "risk premium". Note that a negative value of $\theta_i$ means that taxable bonds have a nonpecuniary disadvantage relative to tax-exempt bonds, so that equalization of risk-adjusted yields requires a deduction of $\theta_i$ from the actual yield on taxable bonds. Thus, the marginal risk premium $\theta_i$ can be treated as analogous to a rate of tax on taxable bond income.

The following equation describes the relationship between taxable and tax-exempt N-period bonds if both are newly issued at par:

$$1 = \sum_{i=1}^{N} \{R[1 - (\tau^*_i - \theta_i)](1 + R_m)^{-i}\} + (1 + R_m)^{-N} \qquad \text{(B1.1)}$$

Note that $\theta_i$ is treated as the addition to the taxable bond yield required to make taxable and tax-exempt bonds equivalent on a risk-adjusted basis. Equation (B1.1) can be used to derive

$$s = \sum_{i=1}^{N} (\tau^*_i - \theta_i) w_i, \quad \text{where}$$

$$w_i = R_m(1 + R_m)^{-i}/[1 - (1 + R_m)^{-N}] \qquad \text{(B1.2)}$$

The left-hand side of (B1.2) is the implicit tax rate, defined as $s = 1 - R_m/R$. The right-hand side is a weighted average of the expected future marginal tax rate less the average marginal risk premium on municipal bonds.

Let $\pi_t$ denote the properly discounted average marginal risk premium. For any investor who holds both tax-exempt and taxable securities, the following relationship holds.

$$s_t = E_t - \pi_t \quad \text{where}$$

$$E_t = \sum_{i=1}^{N} w_i E(\tau_{t+i} | \Omega_t), \qquad \pi_t = \sum_{i=1}^{N} w_i \theta_i, \qquad \text{and}$$

$$w_i = R_m(1 + R_m)^{-i}/[1 - (1 + R_m)^{-N}] \qquad \text{(B1.3)}$$

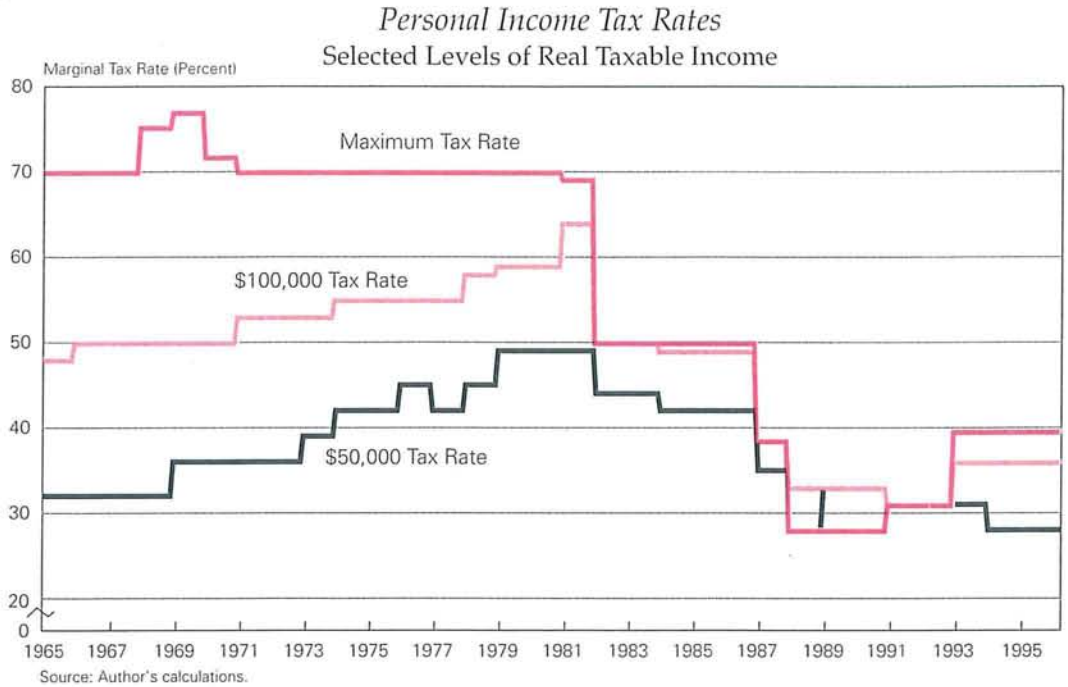
Thus, the spot implicit tax rate is the expected average future marginal tax rate less the average future marginal risk premium. A slight rewriting of (B1.3) tells us that the expected average marginal tax rate must be equal to the *risk-adjusted* implicit tax rate, defined as the implicit tax rate plus the average marginal risk premium on municipal bonds. Thus, in an efficient market the expected future tax rate is the spot implicit tax rate plus a risk premium.

$$E_t = s_t + \pi_t \qquad \text{(B1.4)}$$

President Clinton's Omnibus Budget Reconciliation Act.[6] Note that while the maximum tax rate declined over the 1965–86 period, the marginal tax rates for lower income levels rose. This reflected, in part, the bracket effect of inflation, as rising prices pushed taxpayers into higher income tax brackets even though their real taxable income remained constant. The indexing of tax brackets has mitigated this effect.

[6] This Act also contained a provision to treat market discounts on municipal bonds as ordinary income. Prior to this, market discounts were taxed at the capital gains rate.

Figure 4



Source: Author's calculations.

## *III. Implicit Tax Rates as Forecasts of Future Tax Rates*

A well-known property of optimal forecasts is that the actual value of a variable being forecasted is equal to the value forecast plus a forecast error. Thus, if $E_{t,t-i}$ is the expected or forecasted value of a variable at time $t$, the forecast being made at time $t$-$i$, and if $A_t$ is the realized or actual value of that variable at time $t$, then the following relationship must hold.

$$A_t = E_{t,t-i} + \varepsilon_t \qquad (3)$$

where the term $\varepsilon_t$ is the forecast error. A further property of an optimal forecast is that the forecast error will have a zero mean, and it will be uncorrelated with the forecast value, $E_{t,t-i}$. If the mean forecast error is nonzero, the forecaster is neglecting some systematic information. If the forecast error is correlated with the forecast, relevant information in the form of missing variables is also being neglected. In either case, the forecast will not be optimal.

An example illustrating the use of the relationship between an optimal forecast and the realized value of the variable is Shiller's 1981 study of the efficiency of the U.S. stock market. If, as Efficient Markets Theory suggests, the spot price of a common stock contains all the relevant information about future stock returns, the spot price will be an optimal forecast of the present value of future dividends. Shiller constructed a time series for the realized, or ex post, value of a share of stock using actual dividends paid in the future. Thus, his measure of $A_t$ is the present value of actual dividends paid. Noting that one of the implications of optimal forecasting is that the variance of the realized outcomes must be less than the variance of the forecast, Shiller compared the variance of his series for the "true" value of the S&P 500 (the present value of dividends) with the variance of the spot S&P 500 index. He found that the variance of the spot price exceeded the variance of the present value of dividends. His result can be interpreted in three ways. Either the stock price is not a present value of future dividends, or the market develops inefficient estimates of the true price, or some assumption underlying the analysis is incorrect. Shiller concluded that the U.S. stock market is inefficient.

### *The Model*

Our analysis follows this optimal forecasting approach. It does not suffer from some of the shortcom-

ings of Shiller's study. The bonds we use in constructing ex post tax rates have a finite life, not the perpetual life of common stock. As a result, we can know the path of a representative investor's tax rates throughout the life of the instrument, and we do not need to make assumptions about future values. Also, the coupons on both municipal and Treasury bonds are known at the time the bonds are issued, and they do not adjust according to the (perhaps complicated) dynamics involved in corporate dividend decisions.

However, the tests appropriate for this study cannot be as simple as Shiller's variance bounds test. In an efficient market, the optimal forecast of the average future tax rate over the life of a bond is the *risk-adjusted* implicit tax rate. In the notation of Box 1, the risk-adjusted implicit tax rate is $s_t + \pi_t$, where $s_t$ is the spot implicit tax rate and $\pi_t$ is the risk premium on municipal bonds. Now, if $T_t$ is the ex post tax rate, then optimal forecasting means that $T_t = (s_t + \pi_t) + \varepsilon_t$. Unfortunately, the risk premium is not observable, and it is likely to be correlated with the implicit tax rate and difficult to disentangle from the forecast error. Thus, a simple comparison of the variance of the observed implicit tax rate with the variance of the ex post implicit tax rate will tell us nothing about the information content of the implicit tax rate.

An alternative approach is to directly model the expected future tax rate. Our model is outlined in Box 2. The result is a regression model of the following form.

$$T_t - \tau_t = \alpha_0 + \alpha_1(s_t - \tau_t) + \alpha_2(1/R_t) + \varepsilon_t,$$

$$\text{where} \qquad \varepsilon_t \sim N(0, \sigma^2), \qquad E(\varepsilon_t\, \varepsilon_{t-s}) = 0 \qquad (4)$$

in which the excess of the ex post tax rate over the current tax rate $(T_t - \tau_t)$ is a linear function of the excess of the spot implicit tax rate over the current tax rate $(s_t - \tau_t)$ and of the reciprocal of the yield on taxable bonds. The parameter $\alpha_1$, called the "information parameter," measures the information content of the implicit tax rate: $\alpha_1 = 0$ if there is no content, $\alpha_1 = 1$ when the implicit tax rate is an optimal forecast. The parameter $\alpha_2$ measures the risk premium on municipal bonds: It will be positive when tax-exempt bonds expose the investor to more risk than do taxable bonds of equal maturity.

The risk premium on municipal bonds has several sources. Some elements of the risk premium are embedded in the nature of the instruments and of the markets for them. For example, duration differences, default risk differences, and segmentation of the municipal bond market will give rise to risks faced by all investors. Other sources of the risk premium are specific to the investors. For example, in a progressive tax system the more affluent investors will receive a higher average after-tax rate of return, even though they might earn no unusual after-tax return at the margin. This extra average after-tax return is a sort of investor's surplus, or windfall, which will appear as a component of the risk premium because it represents an excess of the actual tax rate over the implicit tax rate.[7]

*The risk premium on tax-exempt bonds should be directly related to the investor's tax rate, because high-tax-rate investors will receive a greater windfall on their intramarginal investments in tax-exempt bonds.*

If the information parameter and risk premium are constant, our model can be estimated using ordinary least squares. However, both parameters are likely to vary over time. The information parameter should vary because, as Box 2 shows, it is interpreted as reflecting the confidence attached to the implicit tax rate as a predictor of tax policy. That confidence should vary as tax rate policy proposals are proposed, considered, and eventually adopted or rejected. The risk premium will also vary as investors change their assessments of the relative risks of taxable and tax-exempt bonds, and as tax code changes alter the investor's surpluses.

### *The Data*

Our data for bond yields are taken from the Salomon Brothers series for Prime Grade GOs and U.S. Treasury bonds. We have estimated our model

[7] The investor's surplus on municipal bonds arises from the progressivity of the tax code. High-income investors are the most eager purchasers of municipal bonds, but the absorption of the total supply of bonds requires that some lower-income "marginal" investors enter the market. This means that high-income investors receive an after-tax return greater than the amount required to make them hold the bonds they chose to purchase.

## Box 2: A Model for Predicting the Ex Post Tax Rate

Following Trczinka (1982) and Fortune (1988), we assume that, for each investor, portfolio equilibrium is described by the following linear relationship between the yields on municipal and Treasury bonds with equal terms to maturity.

$$R_m = \lambda_{0i} + \lambda_{1i}R \qquad \text{(B2.1)}$$

The parameter $\lambda_{0i}$ is the marginal risk premium required by the *ith* investor to hold his chosen quantity of municipal bonds, and the parameter $\lambda_{1i}$ is the investor's after-tax return on a dollar of taxable income, that is, $\lambda_{1i} = (1 - \tau_i^*)$, where $\tau_i^*$ is the expected marginal income tax rate over the investment horizon. This relationship holds for every investor in municipal bonds since, at the margin, the excess of the after-tax return on municipals over the after-tax return on taxable bonds must be just enough to compensate for the risk associated with holding municipal bonds.

Equation (B2.1) can be used derive the *risk-adjusted* implicit tax rate embedded in relative yields.

$$\tau_i^* = s + \lambda_{0i}/R \qquad \text{(B2.2)}$$

where $s = 1 - R_m/R$ is the observed implicit tax rate, common to all investors.

We assume that the ex post tax rate is a weighted average of several "information variables" that are relevant to forecasting future tax rates. For simplicity, we assume two information variables: the risk-adjusted implicit tax rate, $\tau_i^*$, and the spot (current) tax rate. Denoting an investor's ex post tax rate as $T_{it}$, his current tax rate as $\tau_{it}$, and the forecast error as $\varepsilon_{it}$, our basic model is

$$T_{it} = \alpha_1\tau_i^* + (1 - \alpha_1)\tau_{it} + \varepsilon_{it} \qquad \text{(B2.3)}$$

The parameter $\alpha_1$ is the "information parameter," measuring the information content of the risk-adjusted implicit tax rate. This is the key parameter in our analysis. Substituting (B2.2) into (B2.3) results in the following model whose parameters can be estimated.

$$T_{it} - \tau_{it} = \alpha_1(s_t - \tau_{it}) + \alpha_{2i}(1/R_t) + \varepsilon_{it} \qquad \text{(B2.4)}$$

in which $\alpha_1$ is the information parameter, assumed the same for all investors, and $\alpha_{2i} = \alpha_1\lambda_{0i}$ measures the influence of the risk premium for a specific investor.

Our model assumes that the expected value of the ex post tax rate is a weighted average of the risk-adjusted implicit tax rate and of the current tax rate. The use of a mixed forecast is justified as follows. Let $y$ be a random variable, and $(x_1, x_2)$ be two orthogonal forecasts of $y$. Then $f_1 = y - x_1$ and $f_2 = y - x_2$ are the forecast errors. Assume that they are joint normally distributed, and consider a forecast of y that is a weighted average of the two information variables, $x_1$ and $x_2$. The composite forecast error is $z = y - [\beta x_1 + (1 - \beta)x_2]$, which can be written as $z = \beta f_1 + (1 - \beta)f_2$. The variance of the composite forecast error is

$$\sigma_z^2 = \beta^2\sigma_1^2 + (1 - \beta)^2\sigma_2^2.$$

If investors choose the $\beta$ that minimizes the mean-squared error, that is, minimizes $\sigma_z^2$, the optimal value of $\beta$ is $\beta = \sigma_2^2/(\sigma_2^2 + \text{J}_1^2)$. Thus, the optimal weight on an information variable will be greater, the smaller the total variance explained by that variable, that is, the more "precise" are forecasts using that variable. The value of $\alpha_1$ in (B2.3) is equivalent to this $\beta$.

for 1-year, 5-year, and 10-year terms. Longer terms are excluded because, as we have shown above, the yields are influenced by factors difficult to control for, and the implicit tax rates at long maturities display strange behavior.

Data for the ex post tax rate were calculated using the marginal tax rates in each year for investors having nominal taxable income equivalent to a specific real income level in 1980 dollars. Four marginal tax rate series were constructed: the maximum tax rate, and tax rates for real income levels of $100,000, $75,000, and $50,000.[8] The annual tax rate was assigned to each month in the year, and the ex post tax rate for each of these representative investors was then calculated using equation (2). These calculations assume that interest is paid semiannually, so the tax rate

[8] Between 1987 and 1990 there was a "bubble" in the tax rate schedule: the tax rate schedule was 15 percent, 28 percent, 33 percent, and 28 percent. Thus, the highest tax rate was 33 percent, but this was not the tax rate paid by those at the highest income levels; that tax rate was 28 percent. Our "maximum" tax rate is 28 percent during this period.

is derived as a weighted average of future tax rates at six-month intervals. Because our bond yield and tax rate data are available monthly for the period 1965 to 1995, the regressions for 10-year, 5-year, and 1-year bonds use the sample periods 1965 to 1985, 1965 to 1990, and 1965 to 1994, respectively.

Our basic model explains the excess ex post tax rate as a function of the excess implicit tax rate and the reciprocal of the Treasury bond yield. The latter serves as a variable capturing the risk premium on municipal bonds. We have added a dummy variable for the Tax Reform Act of 1986, defined as zero before 1986:12 and one thereafter. This dummy variable is excluded from the 10-year bond regressions because the dummy variable is always zero during the sample period for those regressions. The reason for this dummy variable is that the Tax Reform Act dramatically altered the structure of tax rates, as well as the incentives of important financial institutions (particularly commercial banks) to hold municipal bonds.

*The information content of implicit tax rates increases dramatically at times of important legislative debates, and has increased over time as tax rate changes have become more frequent.*

We first estimate a constant-coefficients form of our model, in which it is assumed that the parameters do not change over time. We then estimate the model with time-varying coefficients. To estimate the time-varying parameter version of the model, we assume that each of the parameters is a random walk, with the parameter changes arising from random shocks that are normally distributed with zero mean, constant variances, and zero covariances. Thus, the *ith* parameter evolves over time according to the model

$$\alpha_{it} = \alpha_{i,t-1} + u_{it} \quad u_{it} \sim N(0, \delta_i^2) \quad i = 1,2,3,4 \qquad (5)$$

Equations (4) and (5) form a state-space model, in which equation (4) is the measurement equation and equation (5) describes the transition equations. The parameters $\alpha_{it}$ are the state variables. The only parameters to estimate in this model are the transition variances, $\delta_i^2$. These are estimated using the method of maximum likelihood. After maximum likelihood estimation of the four $\delta_i^2$, a Kalman Filter is used to construct the paths of the state variables.[9]

### *The Results*

Table 2 reports the results of estimating the constant-coefficient version of the model.[10] As noted above, the information parameter should be positive and no greater than one. In addition, the risk premium on tax-exempt bonds should be directly related to the investor's tax rate, because high-tax-rate investors will receive a greater windfall on their intramarginal investments in tax-exempt bonds. Table 2 supports these expectations. In all 12 regressions the information parameter ($\alpha_1$, the coefficient for the excess implicit tax rate) is positive and statistically significant. It is roughly the same magnitude, about 0.10, for nearly all terms. This indicates that the implicit tax rate contains statistically significant information about future tax rates, but that the bulk, roughly 90 percent, of the relevant information is in the current tax rate. In short, implicit tax rates have some value in forecasting future tax rates, but not much.

In all but one of the regressions, the effect of the risk premium on municipal bonds, captured by the coefficient $\alpha_2$, is positive, as expected.[11] Also, as expected, the size of $\alpha_2$ is generally related to the investor's taxable income level: It is highest for the maximum tax-rate investor and, in most cases, it declines as taxable income (and tax rates) fall. This is consistent with our knowledge that higher-tax-rate investors receive a larger average after-tax return from investment in municipal bonds. The coefficient on the 1986 Tax Reform Act dummy is positive in all cases, and typically is significant. Thus, after the Tax Reform Act there was an increase in expected future taxes

[9] The time-varying parameter estimation assumes the following initial conditions. The unconditional mean of the four parameters (zero) sets the starting value of the parameters. The initial covariance matrix of the parameters is taken as a diffuse prior; specifically, the value 1,000 is assigned to each initial variance.

[10] Ordinary least squares showed very high serial correlation. To correct for this we used Hannan's method of correction with spectral analysis. The Q statistics reported in Table 2 show that this eliminated serial correlation over a 36-lag period for the 1-year bonds, and for the 5-year bonds with a maximum tax rate. The correction failed for the other regressions.

[11] Box 2 shows that the size of the risk premium is $\lambda_0 = \alpha_2/\alpha_1$. Because $\alpha_1$ is roughly 0.10, $\lambda_0$ is roughly 10 times the value of $\alpha_2$.

Table 2
*Linear Regression Model*
Constant Coefficients

$$T_t - \tau_t = \alpha_0 + \alpha_1(s_t - \tau_t) + \alpha_2(1/R_t) + \alpha_3 TRA86 + \eta_t$$

| Coeff/ t-stat | 1-Year 1965:1 to 1994:12 Income Level | | | | 5-Year 1965:1 to 1990:12 Income Level | | | | 10-Year 1965:1 to 1985:12 Income Level | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Max | $100K | $75K | $50K | Max | $100K | $75K | $50K | Max | $100K | $75K | $50K |
| $\alpha_0$ | −.0047 | −.0074 | −.0046 | −.0020 | −.0104 | −.0050 | −.0043 | −.0027 | −.0200 | −.0067 | −.0110 | −.0012 |
| $t_{\alpha 0}$ | −1.05 | −2.88 | −1.14 | −.39 | −2.97 | −1.55 | −1.57 | −.77 | −3.70 | −1.42 | −2.65 | −.31 |
| $\alpha_1$ | .0968 | .0876 | .0368 | .0732 | .1383 | .1275 | .0574 | .1255 | .1053 | .1058 | .0638 | .0577 |
| $t_{\alpha 1}$ | 5.22 | 6.49 | 4.51 | 5.89 | 5.69 | 6.19 | 5.11 | 6.72 | 3.40 | 4.42 | 3.51 | 3.41 |
| $\alpha_2$ | .1205 | .0524 | .0148 | −.0050 | .2540 | .0819 | .0150 | .0396 | .1505 | .0889 | .1183 | .0475 |
| $t_{\alpha 2}$ | 2.24 | 1.41 | .70 | −.17 | 3.81 | 1.70 | .55 | 1.00 | 1.73 | 1.40 | 2.75 | 1.08 |
| $\alpha_3$ | .0072 | .0582 | .0065 | .0050 | .0694 | .0490 | .0209 | .0160 | n.a. | n.a. | n.a. | n.a. |
| $t_{\alpha 3}$ | .74 | 7.06 | 1.59 | 1.02 | 6.17 | 5.86 | 4.27 | 2.38 | | | | |
| DW | 1.96 | 1.89 | 1.82 | 1.71 | 1.99 | 1.61 | 1.66 | 1.32 | 1.66 | 1.59 | .82 | 1.43 |
| $Q_{36}$ | 11.10 (99%) | 50.66 (5%) | 25.64 (90%) | 35.47 (49%) | 29.84 (76%) | 90.76 (0%) | 168.90 (0%) | 179.20 (0%) | 97.09 (0%) | 167.3 (0%) | 1063.0 (0%) | 219.24 (0%) |

Hannan-Efficient estimation was used. This is a transformation of variables by spectral analysis to correct for serial correlation. The statistic $Q_{36}$ tests for any serial correlation over a 36-month period. The significance level is in parentheses: a significance level of s% means a (1-s)% chance that the observed Q is due to chance.
n.a. not applicable

relative to the information contained in the risk-adjusted implicit tax rate.

Table 3 reports the results of estimating the model in a state-space form, in which the coefficients are time-varying parameters following a random walk. In this model, the parameters, designated $\delta_i$, are the standard errors of the shocks in the transition equations. These are, of course, nonnegative. A very small value of $\delta_i$ means that the associated coefficient, $\alpha_i$, moves very little over time; a large $\delta_i$ means that $\alpha_i$ varies considerably over time.

Our primary interest is in the value of $\delta_1$, which measures the standard deviation of shocks in the information parameter. Shocks to the information parameter are statistically significant for all but one of the 5-year and 10-year cases; the one exception is at the low ($50,000) taxable income level, for which the model is least valid. For three out of the four 1-year cases, $\delta_1$ is not statistically significant. This result is plausible, because one would expect that 1-year tax rate forecasts would be dominated by the current tax rate, and that the weight on the 1-year implicit tax rate would not vary much over time.

A Kalman Filter was used to construct the path of the information parameter implied by the estimates in Table 3. The Kalman Filter is a method of optimal updating of state-variable values as new observations of data emerge. The key equation in a Kalman Filter describes the updating of a state variable as a new observation on the dependent variable arrives. In our notation, the updating equation is $\alpha_t = \alpha_{t,t-1} + g_t \varepsilon_t$, where $\varepsilon_t$ is the current forecast error and $g_t$ is the Kalman Gain, which measures how much the parameter forecast adjusts when a forecast error occurs. The Kalman Gain is inversely related to the variance of the forecast error—when the forecast is very uncertain, the new information is given very little weight in the updating formula. This parameter updating equation says that the optimal current value of a parameter at time $t$ is the value that was forecast at the previous period, $\alpha_{t,t-1}$, plus an adjustment related to the forecast error based on the previous period's information.

Figures 5a and 5b report the paths of the information parameter implied by the Kalman Filter for investors paying the maximum personal income tax rate, and for investors with $100,000 of real taxable income, respectively. These are the taxable income

Table 3

*Maximum Likelihood Estimation of Transition Equation Standard Errors*

State-Space Model Using Kalman Filter

$$T_t - \tau_t = \alpha_{0t} + \alpha_{1t}(s_t - \tau_t) + \alpha_{2t}(1/R_t) + \alpha_{3t}TRA86t_t + \varepsilon_t$$

$$\alpha_{it} = \alpha_{i,t-1} + u_{it}\ (i = 0,1,2,3) \qquad u_t \sim N(0, \Delta^2) \qquad \varepsilon_t \sim N(0, \sigma^2)$$

| Coeff/ | One-Year 1965:1 to 1994:12 Income Level | | | | Five-Year 1965:1 to 1990:12 Income Level | | | | Ten-Year 1965:1 to 1985:12 Income Level | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| t-stat | Max | $100K | $75K | $50K | Max | $100K | $75K | $50K | Max | $100K | $75K | $50K |
| $\delta_0$ | .0025 | .0042 | .0010 | .0000+ | .0526 | .0042 | .0297 | .0010 | .0018 | .0166 | .0003 | .0101 |
| $t_{\delta 0}$ | .59 | .23 | 6.43 | .02 | 14.33 | 37.27 | 12.91 | .48 | 4.61 | 7.27 | .13 | 3.17 |
| $\delta_1$ | .0745 | .0546 | .0190 | .0022 | .0810 | .1018 | .0355 | .1060 | .0590 | .1134 | .1423 | .0120 |
| $t_{\delta 1}$ | 18.70 | 1.30 | .34 | .88 | 8.47 | 13.38 | 20.02 | 9.42 | 8.59 | 6.42 | 13.32 | .33 |
| $\delta_2$ | .0000+ | .0000+ | .0000+ | .0000+ | .0000+ | .0000+ | .0000+ | .0001 | .0000+ | .0002 | .0000+ | .0001 |
| $t_{\delta 2}$ | .03 | .02 | .00+ | .00+ | .23 | .16 | 1.66 | .34 | .01 | 1.17 | .11 | .61 |
| $\delta_3$ | .0037 | .0025 | .0004 | .0000+ | .0999 | .1301 | .0355 | .0005 | na | na | na | na |
| $t_{\delta 3}$ | .42 | .32 | .05 | .00+ | 1.11 | 1.24 | 2.04 | .03 | | | | |
| $\sigma$ | .2948 | .2340 | .1441 | .1942 | .1050 | .1630 | .0362 | .2113 | .2796 | .1097 | .1089 | .1917 |

The parameters $\delta_0$, $\delta_1$, $\delta_2$ and $\delta_3$ are the standard errors of the transition equation steps, defined as square roots of the diagonal elements in $\Delta$. All off-diagonal elements are assumed to be zero. The parameter $\sigma$ is the standard deviation of the measurement equation.

levels for which the model should be the most valid.[12] The information parameter paths show that little weight was attached to implicit tax rates in the 1970s. This was a period of few major tax rate changes, so investors placed most of the weight (about 90 percent) on the current tax rate when forming forecasts of future tax rates.

Beginning in the early 1980s, the information parameter rose dramatically. Two major tax rate restructurings occurred in the 1980s, the Economic Recovery Tax Act of 1981 and the Tax Reform Act of 1986. In both cases the maximum personal income tax rate fell sharply, as did the maximum tax rate at the $100,000 real taxable income level. Figure 5a shows a dramatic jump in the information parameter, from 0.10 to about 0.50. The same pattern is seen in Figure 5b, though the rise in the information parameter began before the passage of ERTA in 1981.

The passage and implementation of the 1986 Tax Reform Act was followed by only small effects on the information parameter. It remained at a high level, indicating the continuation of the high information content in the risk-adjusted implicit tax rate, up to 1990, then (according to the 1-year bond data) began a slow decline, returning to its 1970s level by 1995.

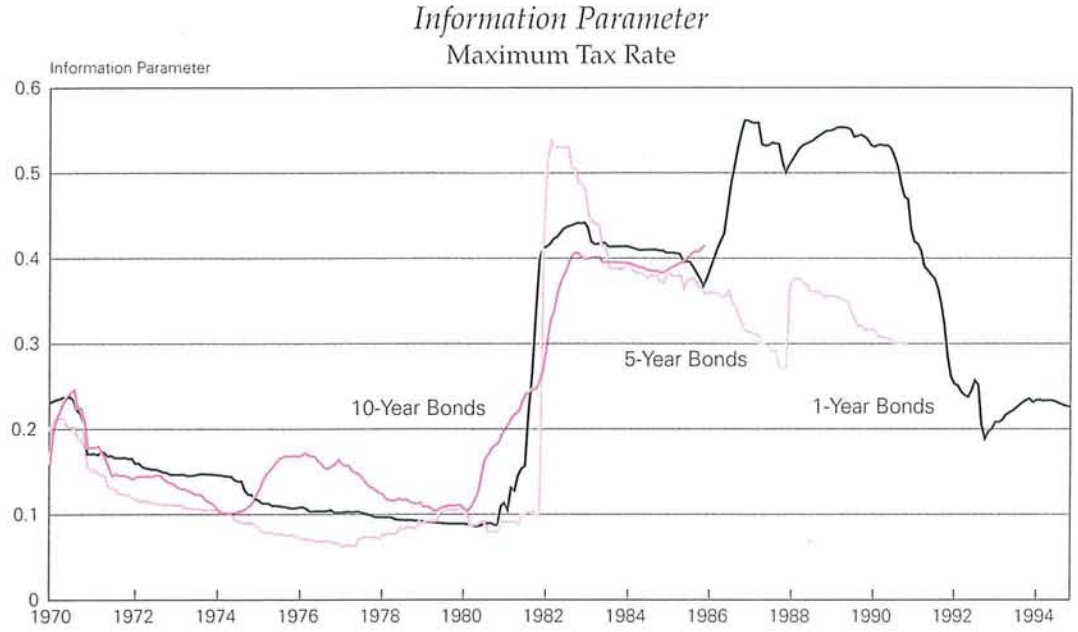
## IV. Summary and Conclusions

The implicit tax rate, derived from yields on Prime Grade general obligations and U.S. Treasury bonds with the same term to maturity, should contain information relevant to the prediction of future tax rates. This study addresses the information content about future tax rates that is embedded in the implicit tax rate. Several specific questions are raised. Does the flat tax debate of 1995–96 explain the high municipal bonds yields (low implicit tax rates) of that period? Is the implicit tax rate determined by the corporate income tax rates, as proposed by Fama and Miller, or by personal income tax rates, as is more commonly thought? Has the information content of the implicit tax rate varied over time?

The first section assesses the previous literature on the relationship between tax policy and relative bond yields. Poterba's event studies of the 1980s are updated, and we find only a weak relationship between implicit tax rates and the proposal, debate, and passage or failure of tax policy legislation that should affect relative bond yields. This section also briefly

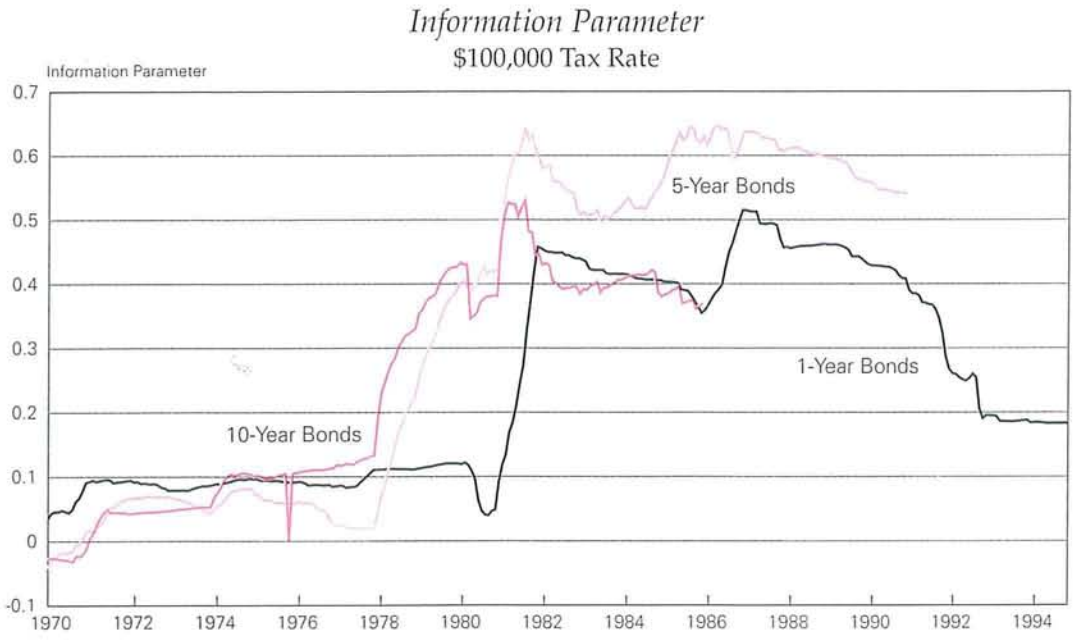
[12] The results are more mixed for investors with $75,000 and $50,000 of taxable income, who would be less likely to conform to our model.

Figure 5a




Source: Author's calculations.

Figure 5b




Source: Author's calculations.

discusses the Fama-Miller hypothesis that only corporate income tax rates affect the relative yield on municipal bonds.

The second section discusses two possible ways of measuring expected future tax rates. The first, the ex ante implicit tax rate, is rejected because previous evidence shows that it can behave in strange ways, especially at long maturities. The second, called the ex post tax rate, represents the actual tax rate that will be paid over the life of a bond by a representative investor.

*We find that the implicit tax rate is a statistically significant predictor of personal income tax rates, and that the information content rose during the period of personal income tax rate variability in the 1980s.*

The third section develops an econometric model of implicit tax rates in which the risk-adjusted implicit tax rate is used as a forecast of the ex post tax rate. The key parameter in this model, called the information parameter, measures the change in the expected future tax rate resulting from a change in the risk-adjusted implicit tax rate. This model is estimated for 1-year, 5-year, and 10-year bonds, over a sample period beginning in January 1965. The model is first estimated as a constant-coefficients regression. Then it is estimated as a state-space model in which the coefficients are the state variables. The latter allows us to use a Kalman Filter to derive the path of the information parameter.

With respect to the narrow question of the flat tax debate as a source of high municipal bond yields (or low implicit tax rates) in 1995–96, the answer appears to be in the negative. Our event analysis shows no statistically significant increase in implicit tax rates during either the July–December 1995 introduction and discussion of H.R. 2060, or the January–February 1996 Presidential primary debate. Unfortunately, the only ex post tax rate series that can be used to test the importance of the flat tax debate (the one-year-ahead series) ends in December of 1994, so our state-space model does not tell us anything about the implicit tax rate debate.

With respect to the broader question of the information content of implicit tax rates, our results are clear. Not only do implicit tax rates contain relevant information, but the information content has changed over time. During the period of tax code stasis in the 1970s, investors appear to have placed a small weight (about 10 pecent) on the implicit tax rate as a predictor of future tax rates, with the primary weight, about 90 percent, put on the current tax rate. During the 1980s, when tax reform was much debated and several significant changes were made in the income tax code, investors recognized the increased probability of tax rate changes by increasing the weight placed on implicit tax rates and reducing their reliance on current tax rates as predictors of tax policy. Thus far, relatively minor changes in tax rates have been made during the 1990s, with the greatest changes being at upper income levels. The weight placed on implicit tax rates has declined back toward its 1970s level. Thus, we find that implicit tax rates are an important predictor around the time of major tax rate changes that are perceived as highly probable. At other times, such as the flat tax debates of 1995–96, implicit tax rates appeared to carry little information about future tax rates.

With respect to the question of whose tax rates matter, our results do not support the Fama-Miller hypothesis. If that hypothesis were valid, we would expect that the implicit tax rate would have information content for predicting corporate income tax rates, not personal income tax rates. We find that the implicit tax rate is a statistically significant predictor of personal income tax rates, and that the information content rose during the period of personal income tax rate variability in the 1980s.

## *References*


Chalmers, John. 1995. "The Relative Yields of Tax-Exempt and Taxable Bonds: Evidence from Municipal Bonds That Are Secured by Treasury Obligations." Ph.D. Dissertation, William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, NY.

Cook, Timothy. 1982. "Determinants of Individual Tax-Exempt Bond Yields: A Survey of the Evidence." Federal Reserve Bank of Richmond *Economic Review*, May/June, pp. 14–38.

Cooley, Thomas and Edward Prescott. 1976. "Estimation in the Presence of Stochastic Parameter Variation." *Econometrica*, vol. 44, no. 1 (January), pp. 167–84.

Fama, Eugene. 1977. "A Pricing Model for the Municipal Bond Market." Unpublished Working Paper, University of Chicago, March.

Fortune, Peter. 1988. "Municipal Bond Yields: Whose Tax Rates Matter?" *National Tax Journal*, vol. 41, no. 2 (June), pp. 219–34.

———. 1991. "The Municipal Bond Market, Part I: Politics, Taxes and Yields." *New England Economic Review*, September/October, pp. 13–36.

Harvey, Andrew. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. New York: Cambridge University Press.

Kochin, Levis and Richard Parks. 1988. "Was the Tax-Exempt Bond Market Inefficient or Were Future Expected Tax Rates Negative?" *Journal of Finance*, vol. 43, no. 4 (September), pp. 913–31.

Miller, Merton. 1977. "Debt and Taxes." *Journal of Finance*, vol. 32 (May), pp. 261–75.

Park, Sangkyun. 1995. "The Relationship Between Government Financial Condition and Expected Tax Rates Reflected in Municipal Bonds." Federal Reserve Bank of New York *Staff Reports*, November.

Poterba, James M. 1986. "Explaining the Yield Spread between Taxable and Tax-Exempt Bonds: The Role of Expected Tax Policy." In Harvey S. Rosen, ed., *Studies in State and Local Public Finance*, pp. 5–49. Chicago, IL: University of Chicago Press.

———. 1989. "Tax Reform and the Market for Tax-Exempt Debt." *Regional Science and Urban Economics*, vol. 19, no. 3 (August), pp. 537–62.

Shiller, Robert. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *The American Economic Review*, vol. 71 (June), pp. 421–36.

Trzcinka, Charles. 1982. "The Pricing of Tax-Exempt Bonds and the Miller Hypothesis." *Journal of Finance*, vol. 37, no. 4 (September), pp. 907–23.

# The Use of Capital Ratios to Trigger Intervention in Problem Banks: Too Little, Too Late

*Joe Peek and Eric S. Rosengren*

*Peek is Professor of Economics, Boston College, and Visiting Economist, Federal Reserve Bank of Boston. Rosengren is Vice President and Economist at the Bank.* *The authors thank Peggy Gilligan and Leo Hsu for providing valuable research assistance. The views expressed are those of the authors, and do not necessarily reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System.*

A wave of depository institution failures and dramatic losses to deposit insurance funds occurred in the 1980s and continued into the early 1990s. In response, the Congress passed a series of bank regulatory acts intended to address the problems that led to the crisis and prevent its recurrence. The Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA) was the capstone of this transformation of banking legislation. Two key provisions of FDICIA were designed specifically to reduce the cost of troubled banks to the deposit insurance fund—early closure of failing institutions and early supervisory intervention in problem banks, the latter referred to as prompt corrective action.

The prompt corrective action (PCA) legislation was intended to supplement existing supervisory authority and make intervention by bank supervisors both more timely and less discretionary. To limit supervisory forbearance, delay or failure to take appropriate action at financially troubled banks, the legislation requires mandatory action by bank regulators at those problem banks that potentially could risk deposit insurance funds. One important requirement for effective early intervention is the ability to identify problem banks early enough to give supervisory intervention an opportunity to affect bank behavior before it is "too late."

A second requirement is to act on that identification in a timely manner. By legislating mandatory intervention, PCA is intended to fill the gaps in the preexisting regulatory intervention framework. In order to minimize the potential for forbearance, mandatory intervention should occur at the time regulators determine that a bank fits the profile of a problem bank with a high probability of failure, putting the deposit insurance fund at risk. Any delay in intervening, once such a problem bank has been identified, is a form of forbearance. Thus, a critical component of PCA is the trigger that initiates mandatory actions by regulators, reported capital ratios. Rules based on such thresholds can be

simple to follow and politically expedient, but several recent studies (for example, Jones and King 1992, 1995) have argued that capital ratios frequently are a lagging indicator of banking problems and often fail to identify severely troubled institutions in a timely manner. If capital ratios do not identify problem institutions early, then prompt corrective action will fall short of its potential for protecting the deposit insurance fund.

This article considers whether the capital ratio thresholds that trigger PCA intervention provide sufficient lead time for successful intervention at troubled banks. The study finds that because PCA is based on a lagging indicator, it is likely to trigger intervention in problem banks only after they have been identified by examiners. Because they rely on far more information than the capital ratio, examiners often are aware of problems well before a bank becomes "undercapitalized" as defined by the PCA provisions contained in FDICIA. Since PCA is intended to supplement actions already taken by examiners, it is reassuring that examiners usually identify problems before PCA guidelines are triggered. Ideally, supervisory intervention will occur early enough to avoid more significant problems. By intervening early to alter bank behavior, it is hoped that failures can be prevented, minimizing the need for FDIC assistance, and if not, that failure results in the minimum cost to the deposit insurance fund.

*The prompt corrective action legislation was intended to supplement existing supervisory authority and make intervention by bank supervisors both more timely and less discretionary.*

Requiring reported data to more accurately reflect a bank's financial health would improve rules that use reported capital ratios to trigger intervention in problem institutions. Still, no simple capital rule can substitute for examiner judgment based on more comprehensive information. Nonetheless, we find that raising the PCA capital ratio thresholds that identify undercapitalized banks would enable the intervention triggered by PCA to mimic more closely the timing of examiner identification of problem banks.

The first section of the article discusses early intervention in problem banks under PCA. The second section examines the identification of problem banks using examiner ratings. The third section considers alternative triggers and shows that triggers based on CAMEL rating downgrades would be superior to the current PCA capital ratio thresholds for mandatory intervention in problem banks. The fourth section summarizes the policy implications of the findings and the final section presents conclusions.

## *I. Identification of Problem Institutions under PCA*

To be effective, supervisory intervention must occur while enough time remains to alter a problem bank's behavior. Not only must the signal that triggers the need for intervention identify problem banks accurately, it must also be timely. It is easy to identify a problem bank at the time of its failure. The challenge is to identify a problem bank in time to prevent its failure or at least in time to alter its behavior in order to limit the losses to the deposit insurance fund. Thus, an appropriate slogan for early intervention might be "the earlier the better." However, such an approach must be tempered by giving appropriate weight to the costs associated with supervisory intervention in banks that are incorrectly identified as "troubled." Earlier identification of potential problem banks will likely be associated with a larger number of non-troubled banks inadvertently being identified as troubled, and the selection of an appropriate trigger for supervisory intervention must recognize this trade-off.

Prompt corrective action as defined in FDICIA identifies a problem bank in need of intervention on the basis of its reported capital ratio.[1] Table 1 highlights the major features of the prompt corrective action provisions. Banks are classified according to their capital adequacy. Banks with total risk-based

[1] Capital thresholds are stated in the prompt corrective action provisions of FDICIA in terms of both leverage ratios and risk-based capital ratios. This study focuses only on leverage ratio thresholds. First, risk-based capital ratios are not available before 1990. Second, for the period in New England under study here, a 6 percent leverage ratio, as generally mandated in formal regulatory actions for troubled institutions, tended to be the binding constraint on capital-constrained banks, rather than the risk-based capital ratios. This is consistent with evidence on nationwide samples that leverage ratios and not risk-based capital ratios affected bank behavior (for example, Hancock and Wilcox 1994).

Table 1

*Prompt Corrective Action*

This table highlights the major features of prompt corrective action contained in FDICIA. Banks are classified according to their capital adequacy. Only major restrictions are described; other restrictions and discretionary actions by supervisors can be applied.

| Categories | Capital Thresholds[a] | Major Restrictions |
|---|---|---|
| 1. Well Capitalized | RBC≥10% and LR ≥ 5% | None |
| 2. Adequately Capitalized | RBC ≥ 8% and LR ≥ 4% | None |
| 3. Undercapitalized | RBC < 8% or LR < 4% | 1. Capital restoration plan<br>2. Suspend dividends<br>3. Restrict asset growth<br>4. Prior approval for expansion |
| 4. Significantly Undercapitalized | RBC < 6% or LR < 3% | 1. Require recapitalization<br>2. Restrict transactions with affiliates<br>3. Restrict interest rates paid<br>4. Further restrict asset growth<br>5. Prohibit deposits from correspondents<br>6. Hire, replace senior management |
| 5. Critically Undercapitalized | $\frac{\text{tangible equity}}{\text{total assets}} \leq 2\%$ | Receivership or conservatorship within 90 days unless exempted by primary regulator and FDIC |

[a]RBC = risk-based capital ratio: total capital, including equity, subordinated debt, and preferred stock, divided by risk-weighted assets.

LR = leverage ratio: tier 1 capital, including equity capital, divided by total average assets.

capital ratios of 10 percent or more and (tier 1) leverage ratios equal to or above 5 percent are considered well capitalized and have no significant mandatory restrictions on activities. Banks with total risk-based capital ratios of 8 percent or higher and leverage ratios of 4 percent and above are rated adequately capitalized and have no substantial mandatory restrictions on activities. Banks with risk-based capital ratios below 8 percent or leverage ratios below 4 percent are considered undercapitalized; they have restrictions on dividends and on asset growth, and they must provide a capital restoration plan. Banks with risk-based capital ratios below 6 percent or leverage ratios below 3 percent are rated significantly undercapitalized and, in addition to the restrictions for undercapitalized banks, face further restrictions on asset growth and on interest rates paid. They also must recapitalize and senior management may be replaced.

PCA is likely to be most effective if the specific capital targets are set so as to trigger intervention that is sufficiently prompt to reduce the costs to the Federal Deposit Insurance Corporation (FDIC). However, a number of studies have indicated that the leverage ratio can better be described as a lagging indicator of a bank's financial health, rather than a leading or even coincident one. (See, for example, Jones and King 1992, 1995.) Because capital is a lagging indicator and examiners utilize information beyond capital ratios when taking supervisory action, the PCA triggers rarely identify problem institutions before examiners.

This appears to be an accurate characterization of the period surrounding the recent banking crisis in New England. Figure 1 shows, for the period from the first quarter of 1988 to the fourth quarter of 1994, the share of New England banks with nonperforming loans (loans more than 90 days past due plus nonaccruing loans) in excess of 5 percent of their assets, the share of banks with a leverage ratio below 4 percent, and the number of New England bank failures. Clearly, New England banks began having serious problems with their loan portfolios well before those problems were reflected in their capital ratios.

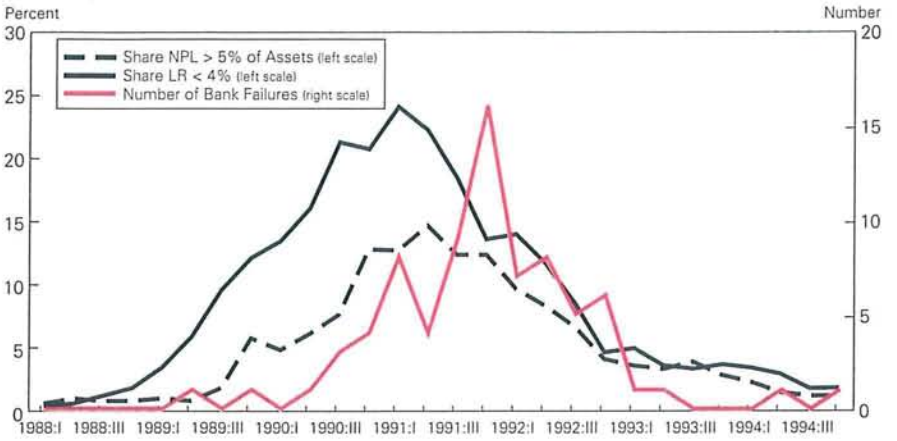
In part, this is because reductions in reported leverage ratios associated with the deterioration in a bank's health are often delayed until the bank undergoes a supervisory examination. Banks tend to be slow to provision for possible loan losses (which, other things equal, reduces their capital), often adding to loan loss reserves only after the problems have been identified, rather than in anticipation of problems. In fact, reported leverage ratios for troubled banks that have not been subjected to a supervisory examination often overstate the institution's financial health. Consequently, many banks experience a large reduction in their reported capital ratios as a consequence of requirements to charge off loans and replenish loan loss reserves that are imposed after an examination (Peek and Rosengren 1996).

Thus, it should not be surprising to find, as shown in Figure 1, that reductions in leverage ratios below

the 4 percent threshold tended to lag behind the deterioration in bank health. But, perhaps more important, the share of banks with a leverage ratio below the 4 percent "undercapitalized" threshold barely leads the wave of bank failures. This is particularly troubling insofar as the official failure date of a bank often occurs well after its economic failure. Regulators often delay official closure after the decision has been made to allow the FDIC time to arrange the disposition of the failed bank.

Figure 1

*Early Warning Indicators for the New England Banking Crisis*




It should also be noted that the number of failures may substantially understate the magnitude of the banking crisis. In many cases, as the health of individual banks deteriorates, multibank holding companies choose, or are induced by regulators, to consolidate banks within their holding company. For example, Bank of New England Corporation had 11 commercial and savings bank subsidiaries in the first quarter of 1988. By the time of its failure in January 1991, the holding company had, through affiliate mergers, reduced the number of its subsidiaries from 11 to only three. Thus, the three failures of Bank of New England Corporation subsidiaries recorded in 1991:I in a sense understate by eight the number of bank failures.

## *II. Identification of Problem Banks through CAMEL Ratings*

Supervisory ratings of banks are another means of identifying problem banks, although the ratings are not made public. Bank supervisors rate the financial condition of a bank considering the capital adequacy, asset quality, management quality, earnings potential, and liquidity of the institution (CAMEL). Each component is evaluated on a scale from 1 to 5, with 1 the highest rating and 5 the lowest. The composite CAMEL rating, which also ranges from 1 to 5, provides an assessment by examiners of the overall strength of a banking institution.[2] Banks with a composite rating of 1 (sound in every respect, flawless performance) and 2 (fundamentally sound, only minor correctable weaknesses in performance) are resistant to external economic and financial disturbances and are not likely to be constrained by regulatory oversight. As a bank's composite rating falls to 3 (remote probability of failure, flawed performance), 4 (potential of failure, performance could impair viability), or 5 (high probability of failure, critically deficient performance), the supervisor's assessment of the likelihood of failure increases.

While capital is a critical component of CAMEL ratings, other elements also play a significant role. The capital position of the bank is considered in relation to the riskiness of its assets, the extent of problem assets in the bank, the severity of its problems with past due loans, the bank's policies and controls for evaluating and monitoring interest rate and credit risk, the availability of earnings to meet the needs for growth and to cover expected losses, and the ability to meet liquidity needs based on the stability and maturity of the assets and liabilities of the bank. Even the C in the CAMEL rating considers more than the regulatory

[2] The composite CAMEL rating and each of the components are always evaluated at the end of a comprehensive examination. In addition, the composite CAMEL rating may be changed between examinations if off-site monitoring indicates a significant change in the financial condition of the bank.

capital standards; for example, capital is considered relative to the size and type of institution.

A large number of FDIC-insured New England savings and commercial banks fell below the 4 percent leverage ratio threshold during the period from the first quarter of 1988 through the fourth quarter of 1994. Table 2 shows the CAMEL ratings at those banks at the time they fell below the 4 percent leverage ratio threshold, consistent with entering the PCA undercapitalized category. By the time banks became undercapitalized by this definition, nearly all had been previously recognized by supervisors as having serious problems. Of the 136 institutions that fell below the 4 percent leverage ratio threshold, 76 percent had CAMEL ratings of 4 or 5, indicating a high probability of failure. Even so, the CAMEL ratings were not perfect indicators, for 10 banks had a CAMEL rating of 1 or 2 at the time they fell below the 4 percent threshold.[3]

Table 3 shows the leverage ratio at the time of a CAMEL downgrade for all FDIC-insured New England savings and commercial banks that were downgraded to a rating of 3, 4, or 5 during the period from the first quarter of 1988 through the fourth quarter of 1994. Banks that get downgraded to a CAMEL rating of 3 or 4 tend to be adequately or well capitalized under the PCA guidelines. Particularly striking is the fact that only 15 of the 182 banks downgraded to a CAMEL 4 rating had fallen below the 4 percent leverage ratio threshold. Thus, most banks classified as having a "possibility of failure" would not have been identified as problem banks using the PCA capital thresholds.

Not only were more than two-thirds of the banks downgraded to a CAMEL 4 rating considered well capitalized at the time of the downgrade, but 31 of the banks had leverage ratios equal to or exceeding 8 percent, twice the threshold for undercapitalized banks. Even in the case of banks with a CAMEL 5 rating, representing a "high probability of failure," the problems were not fully reflected in reported capital ratios. Of the 115 banks downgraded to a CAMEL 5 rating, 27 had leverage ratios exceeding 4 percent. However, by the time supervisors classified banks as having a high probability of failure, most were at least undercapitalized by PCA standards.

Fortunately, supervisory intervention also occurs independent of PCA triggers. Both informal and formal regulatory actions have been imposed by regulators on many banks, particularly in recent years in New England (Peek and Rosengren 1995, 1996). Furthermore, Peek and Rosengren (1996) find that bank

Table 2
*CAMEL Ratings at Banks When Crossing 4 Percent Leverage Ratio*
The sample of banks included in this table are all FDIC-insured New England savings and commercial banks that fell below the 4 percent leverage ratio threshold between 1988:I and 1994:IV.

| CAMEL Rating | Number of Banks |
|---|---|
| 1 | 1 |
| 2 | 9 |
| 3 | 23 |
| 4 | 57 |
| 5 | 46 |

Table 3
*Leverage Ratio at Time of CAMEL Downgrade*

| | Number of N.E. Banks Where CAMEL Downgraded to:[a] | | |
|---|---|---|---|
| Leverage Ratio | 3 | 4 | 5 |
| Less than 2.0 | 1 | 4 | 43 |
| 2.0–2.5 | 0 | 1 | 10 |
| 2.5–3.0 | 1 | 4 | 14 |
| 3.0–3.5 | 0 | 2 | 14 |
| 3.5–4.0 | 2 | 4 | 7 |
| 4.0–4.5 | 6 | 19 | 8 |
| 4.5–5.0 | 11 | 23 | 6 |
| 5.0–5.5 | 11 | 21 | 1 |
| 5.5–6.0 | 17 | 23 | 2 |
| 6.0–6.5 | 31 | 15 | 6 |
| 6.5–7.0 | 29 | 14 | 2 |
| 7.0–7.5 | 24 | 13 | 0 |
| 7.5–8.0 | 17 | 8 | 0 |
| Greater than or equal to 8.0 | 86 | 31 | 2 |
| Total | 236 | 182 | 115 |

[a]CAMEL ratings include interim changes between exams.

[3] The 10 banks with a CAMEL rating of either 1 or 2 at the time they crossed the 4 percent leverage ratio threshold appear puzzling at first glance. However, this pattern can be explained in large part by the infrequent examinations at many banks prior to FDICIA, which mandated annual examinations at banks. All 10 instances occurred before the end of the second quarter of 1991. The elapsed time between the most recent examination and the crossing of the leverage ratio threshold at these banks ranged from five to 11 quarters. In the seven cases where a subsequent examination occurred prior to the bank's failure or acquisition, two banks were downgraded to a CAMEL 4 rating and five were downgraded to a CAMEL 5 rating.

Table 4

*Timing of CAMEL Downgrades Relative to Crossing 4 Percent Leverage Ratio*

All FDIC-insured commercial and savings banks in New England are classified as crossing a 4 percent leverage ratio or never dropping below a 4 percent leverage ratio during the 1988:I to 1994:IV period. For banks that cross a 4 percent leverage ratio, spans of eight quarters before and after crossing the threshold are examined. Negative numbers indicate the number of quarters that the CAMEL downgrade preceded the leverage ratio falling below 4 percent, while positive numbers indicate the number of quarters the CAMEL downgrade followed the leverage ratio falling below 4 percent.

| Quarters until Downgrade, Relative to Crossing 4% Leverage Ratio | CAMEL Rating | |
|---|---|---|
| | 4[a] | 5 |
| −8 | 1 | 0 |
| −7 | 2 | 0 |
| −6 | 3 | 0 |
| −5 | 9 | 1 |
| −4 | 11 | 2 |
| −3 | 12 | 0 |
| −2 | 16 | 4 |
| −1 | 17 | 8 |
| 0 | 25 | 30 |
| 1 | 15 | 28 |
| 2 | 5 | 11 |
| 3 | 4 | 10 |
| 4 | 0 | 8 |
| 5 | 0 | 1 |
| 6 | 0 | 1 |
| 7 | 0 | 0 |
| 8 | 0 | 0 |
| Already at that rating | 6 | 1 |
| Already at lower rating | 1 | NA |
| Failed or acquired before reaching rating | 8 | 10 |
| Remained above rating | 1 | 21 |
| Never dropped below 4% LR | 399 | |
| But go as low as CAMEL 4 | 99 | |
| But go as low as CAMEL 5 | 9 | |

[a]Banks that skipped a 4 rating because they were downgraded from 3 or higher to 5 in a single quarter are included in this column, as well as in the next column.

supervisors do tend to implement formal actions prior to banks becoming undercapitalized as defined by PCA provisions. This discrepancy between the PCA thresholds and the capital ratio at which most supervisory intervention occurs suggests that examiners do not view capital alone as a sufficient statistic for bank health.

For New England commercial and savings banks, Table 4 shows the timing of CAMEL downgrades relative to crossing the 4 percent leverage ratio threshold. The analysis considers a window that spans eight quarters before and eight quarters after a bank crosses the 4 percent leverage ratio threshold. Negative numbers indicate the number of quarters that the CAMEL downgrade preceded the leverage ratio falling below 4 percent, while positive numbers indicate the number of quarters that the CAMEL downgrade followed the leverage ratio falling below 4 percent. Of the 136 New England banks that fell below the 4 percent leverage ratio threshold, 78 (57 percent) of the banks (including seven already at that rating or below more than eight quarters prior to the downgrade) were downgraded to a CAMEL 4 rating before crossing the capital threshold and 25 banks (18 percent) crossed the threshold in the same quarter.[4]

*With the implementation of the FDICIA requirement of annual supervisory examinations, CAMEL ratings will better reflect the current health of banks.*

Many of the downgrades occurred substantially before the reported loss of capital, with 33 banks (including the seven already at that rating or below more than eight quarters prior to the downgrade) downgraded as long as one year prior to crossing the 4 percent leverage ratio threshold. Only 24 banks (18 percent) were downgraded after crossing the capital threshold, and all of these downgrades occurred within three quarters of crossing the 4 percent leverage ratio threshold. On the other hand, many banks were viewed as problems by supervisors even though they never became undercapitalized, as defined by the 4 percent leverage ratio threshold. In fact, 99 banks that received a CAMEL 4 (or lower) rating did not fall below the 4 percent leverage ratio threshold during the sample period.

Banks are generally downgraded to a CAMEL 5 rating only after becoming undercapitalized. However, 16 banks were downgraded to a CAMEL 5 before

[4] Banks are clustered concurrently with the downgrade because many banks are forced to reserve for loan losses (which reduces capital) as a consequence of the examination that resulted in the downgrade.

becoming undercapitalized, with another 30 banks downgraded in the same quarter that they crossed the threshold. Since a CAMEL 5 rating indicates a high probability of failure, these are the banks where the mandatory PCA restrictions should already have been in place in order to minimize the risks and costs of failure, as was intended in the PCA provisions. Nine additional banks were downgraded to a CAMEL 5 rating without ever crossing the 4 percent leverage ratio.

With the implementation of the FDICIA requirement of annual supervisory examinations, CAMEL downgrades should perform even better than in the period prior to FDICIA. With more frequent exams, CAMEL ratings will better reflect the current health of banks. Thus, as a bank's financial health deteriorates, CAMEL downgrades will tend to occur sooner. The more frequent exams should also mitigate the extent to which reported capital ratios diverge from values that accurately reflect the bank's financial health.

## *III. The Timing of Alternative Triggers*

The regulatory intervention component of FDICIA included two changes designed to reduce the cost of troubled banks to the deposit insurance fund. The first was early closure of banks that could not be turned around, before management could undertake second bets resulting in even larger losses. The second was early intervention, changing banks' behavior early enough to ensure that they did not fail or, if they did fail, that failure resulted in minimum cost to the deposit insurance fund. Preventing failure is obviously preferable, because it avoids using any FDIC funds. However, it requires identification and intervention early enough to change bank behavior in a way that will avoid failure.

Table 5 shows the differences in the timing by which banks that eventually failed would have been identified, using five alternative criteria.[5] We focus on failed banks to confirm that timing differences occurred for banks that were clearly troubled.[6] Since more banks are identified as problem banks with CAMEL ratings than under PCA guidelines, we do not want timing differences to be attributed to possible misclassification of healthy banks as troubled. Ideally, the identification would have resulted in actions that prevented failure altogether.

Table 5
*Timing of Identification of Failed Banks*
For the 77 New England commercial and savings banks that failed between 1988:I and 1994:IV, the timing of crossing various thresholds is shown.

| Quarters Identified before Failure | CAMEL Rating = 4 | CAMEL Rating = 5 | Leverage Ratio = 5 | Leverage Ratio = 4 | Leverage Ratio = 3 |
|---|---|---|---|---|---|
| 1–4 | 17 | 44 | 22 | 35 | 47 |
| 5–8 | 44 | 29 | 44 | 35 | 25 |
| 9–12 | 7 | 1 | 9 | 4 | 1 |
| >12 | 7 | 1 | 2 | 1 | 1 |
| Total | 75 | 75 | 77 | 75 | 74 |

CAMEL 4 and CAMEL 5 downgrades occurred at 75 of the 77 failed banks. The two exceptions were small banks that had not been examined for nearly two years prior to their failure. Each of the 77 failed banks crossed the 5 percent leverage ratio threshold that delineates "well capitalized" banks from "adequately" capitalized banks. Seventy-five of them also crossed the PCA 4 percent threshold to become "undercapitalized." The two exceptions were members of failed multibank holding companies with much more poorly capitalized affiliates. One additional failed bank was not identified by the 3 percent leverage ratio threshold, a small bank whose leverage ratio had fallen only slightly below 4 percent by the time of its last call report.

While the total number of failed banks identified differs little across these five alternative thresholds, the relative timing of the identification does differ, especially between the two periods one to four quarters and five to eight quarters prior to failure. The CAMEL 4 threshold dominates in terms of identifying

[5] The leverage threshold crossings analyzed in Table 5 are based only on data reported in the quarterly call reports. Thus, if a bank's capital declines between the time it files its last call report and the time it fails, perhaps because of examiner requirements that the bank substantially increase its loan loss reserves, that decline will not be reflected in this table. Similarly, if a bank's financial health deteriorates rapidly or if the bank is examined infrequently, the failure can occur before an examination that would reflect the weaker condition of the bank, so that the bank's CAMEL rating never reflects the poor condition of the bank. These factors likely account for the fact that the total number of failed banks that cross the 3 percent and 4 percent leverage ratio thresholds or receive a CAMEL 4 or 5 rating is less than the total number of New England banks (77) that failed during this period.

[6] This study is intended to examine procedures to identify problem banks and prevent failure. While PCA also has early closure provisions, effective early intervention should minimize the need for early closure.

failed banks more than three years prior to failure, with seven banks so identified. Only the 5 percent leverage ratio threshold identifies as many as two. The CAMEL 4 threshold again is the winner for more than two years, with a total of 14 banks. Again, the 5 percent leverage ratio is the runner-up, with 11 banks identified. The PCA leverage ratio threshold of 4 percent for undercapitalized banks manages to identify only five banks more than two years prior to failure.

Each of the alternative thresholds identifies a substantial number of failed banks five to eight quarters prior to their failure. However, the sum of banks identified more than a year prior to failure is still dominated by CAMEL 4 downgrades, with 58 banks identified, and the 5 percent leverage ratio threshold, with 55 banks identified. The PCA 4 percent threshold comes in a distant third, with 40 banks identified. Of the 75 banks that crossed the 4 percent leverage ratio threshold, the minimum capital ratio for supposedly adequately capitalized banks, 35 failed within one year of crossing the threshold. This compares to only 17 of the 75 institutions downgraded to a CAMEL 4 rating that failed within one year of the downgrade.

If one purpose of PCA intervention is to prevent forbearance by regulators that allows second bets at troubled institutions, both a downgrade to a CAMEL 4 rating and the crossing of a 5 percent leverage ratio threshold would seem more appropriate thresholds than the 4 percent leverage ratio now specified for the corrective actions contained in the PCA provisions of FDICIA. Both thresholds identify problem banks much earlier than a 4 percent leverage ratio threshold, with the CAMEL 4 threshold having a slight advantage over the 5 percent leverage ratio. For many banks, falling below the 4 percent leverage ratio threshold, now the minimum capital ratio for a bank to be deemed adequately capitalized, has been quickly followed by failure, before any serious attempt could be made to alter bank behavior. Most of the PCA early intervention restrictions apply only after the bank has fallen below the 4 percent threshold, and supervisory intervention based solely on PCA provisions likely would be too late to prevent failure.

However, raising leverage ratio thresholds for defining problem banks can be costly if too many banks are identified as problem banks. Examiner resources are limited, and it is important to concentrate on banks with high probabilities of failure. Also, to the extent that remedial actions are costly for the bank, such actions should not be imposed on banks with little likelihood of failure. Unfortunately, it is difficult to identify the number of banks that posed little problem of failure and were falsely identified as problem banks. While 224 banks (of which 75 subsequently failed) in our sample fell to a rating of CAMEL 4, only 183 banks (of which 77 subsequently failed) fell below a leverage ratio of 5 percent. (157 banks received both a CAMEL rating of 4 and a leverage ratio below 5 percent.) It may be that successful intervention by examiners prevented some banks' capital from falling further. While identifying too many banks as problems has a cost and is a potential risk in introducing higher thresholds, we do not have an estimate for the size of these costs.

## *IV. Policy Implications*

Prompt corrective action can best reduce regulatory forbearance by causing regulators to intervene early in problem banks. Examiners evaluate asset quality, management, earnings potential, and liquidity in addition to capital when they identify banks with a high probability of failure. Since examiners use far more information than capital ratios to identify problem banks, it should not be surprising that examiner rating downgrades tend to identify problem banks earlier than the PCA capital ratio thresholds as currently stated. Thus, earlier PCA intervention in problem banks could be achieved by altering the PCA triggers to make them more closely mimic the timing of problem bank identification by examiners.

*Because examiners want actions to occur early enough to alter behavior to prevent failure, proposals to raise PCA capital ratio triggers, or to set them based on CAMEL ratings, would still serve only as a supplement to existing examiner actions.*

One possibility is to use CAMEL rating down-

grades, rather than capital ratios, as the trigger for PCA intervention. Reported capital ratios are lagging indicators of bank health, in part because some banks have not fully reflected likely future losses in their loan loss reserve. CAMEL ratings do not suffer from this drawback, since they include an evaluation of the bank's capital ratio as well as an evaluation of other factors that examiners believe indicate whether additional problems exist.

In fact, evidence suggests that examiners do have information that is superior to publicly available data such as reported capital ratios, loan loss reserves, and nonperforming loans. For example, a number of studies have found that CAMEL ratings provide information about troubled banks that is not publicly available (Berger and Davies 1994; Gilbert 1993; Gilbert and Park 1994). This should not be surprising, since banks are a repository of private information about their loan customers (for example, James 1987) and this information can be evaluated only by examining a bank's loan files. Furthermore, given that bank management has an incentive to disclose to the public positive rather than adverse information about bank operations, the informational advantage should be even greater for problem banks. Consistent with this hypothesis, Berger and Davies (1994) have found that examination downgrades reveal unfavorable private information about a bank's condition.

A second possibility for earlier PCA intervention is to raise the leverage ratio threshold that triggers the mandatory PCA intervention, for example, to 5 percent. The current 4 percent trigger appears to be too low, resulting in mandatory intervention that would tend to occur only well after the identification of a problem bank by examiners, substantially limiting the potential for PCA legislation to mitigate any possible supervisory forbearance. Such a change is attractive because it retains the characteristic of being a simple rule that is based on a measure that is publicly available and thus verifiable.

Either of these simple proposals would lower the hurdle for intervening in banks.[7] Because examiners want actions to occur early enough to alter behavior to prevent failure, these proposals would still serve only as a supplement to existing examiner actions. Ideally, PCA would not be binding, because examiners would take action well before capital ratios reached even these higher minimum standards.

## V. Conclusions

The prompt corrective action provisions of FDICIA are intended to prevent supervisory forbearance by requiring mandatory intervention in problem banks that have an incentive to take second bets. This legislation is intended to supplement rather than replace the preexisting supervisory intervention framework that includes informal and formal regulatory actions. However, should supervisory forbearance occur, the PCA triggers as currently formulated are likely to ensure that the mandatory intervention requirements will fall far short of their potential. Because the capital ratios that are used as the PCA triggers are lagging indicators of a bank's financial health, the intervention will tend to occur well after examiners identify problems at a bank and, in many cases, leave little time for the intervention to affect bank behavior before it is too late. Of the banks that cross the PCA 4 percent leverage ratio threshold that triggers significant restrictions on their behavior, many fail subsequently, and the failure frequently occurs within one year of crossing the threshold.

PCA triggers that more closely mimic the timing of problem bank identification by examiners would result in more timely intervention in problem institutions. Examiners use far more information than just capital ratios to determine a bank's likelihood of failure. Setting PCA triggers based on CAMEL ratings, or raising the PCA capital ratio triggers, are possible changes that could lead to earlier PCA intervention in problem institutions, encouraging preemptive action that would avoid bank failures requiring deposit insurance funds.

[7] Other, more complicated proposals also have been suggested. For example, Jones and King (1995) suggest adjusting reported risk-based capital ratios to make them more accurately reflect a bank's financial health, by using information on classified loans to adjust the allowance for loan and lease losses (and hence capital) and by raising the risk weights for classified assets. However, even the simple proposals suggested here can provide an improvement over the current PCA triggers.

## *References*


Berger, Allen N. and Sally M. Davies. 1994. "The Information Content of Bank Examinations." Federal Reserve Board Working Paper 94-20.

Gilbert, R. Alton. 1993. "Implications of Annual Examinations for the Bank Insurance Fund." Federal Reserve Bank of St. Louis *Review*, January/February, pp. 35–52.

Gilbert, R. Alton and Sangkyun Park. 1994. "Value of Early Warning Models in Bank Supervision." Manuscript.

Hancock, Diana and James A. Wilcox. 1994. "Bank Capital and the Credit Crunch: The Roles of Risk-Weighted and Unweighted Capital Regulations." *Journal of the American Real Estate and Urban Economics Association*, March, pp. 59–93.

James, Christopher. 1987. "Some Evidence on the Uniqueness of Bank Loans." *Journal of Financial Economics*, vol. 19, pp. 217–35.

Jones, David S. and Kathleen Kuester King. 1992. "The Implementation of Prompt Corrective Action." In *Credit Markets in Transition*, Proceedings of the 28th Annual Conference on Bank Structure and Competition, Federal Reserve Bank of Chicago, May, pp. 68–100.

———. 1995. "The Implementation of Prompt Corrective Action: An Assessment." *Journal of Banking and Finance*, vol. 19, no. 3-4, pp. 491–510.

Peek, Joe and Eric S. Rosengren. 1995. "Bank Regulation and the Credit Crunch." *Journal of Banking and Finance*, vol. 19, no. 3-4, pp. 679–92.

———. 1996. "Will Legislated Early Intervention Prevent the Next Banking Crisis?" Manuscript, Federal Reserve Bank of Boston, April.

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