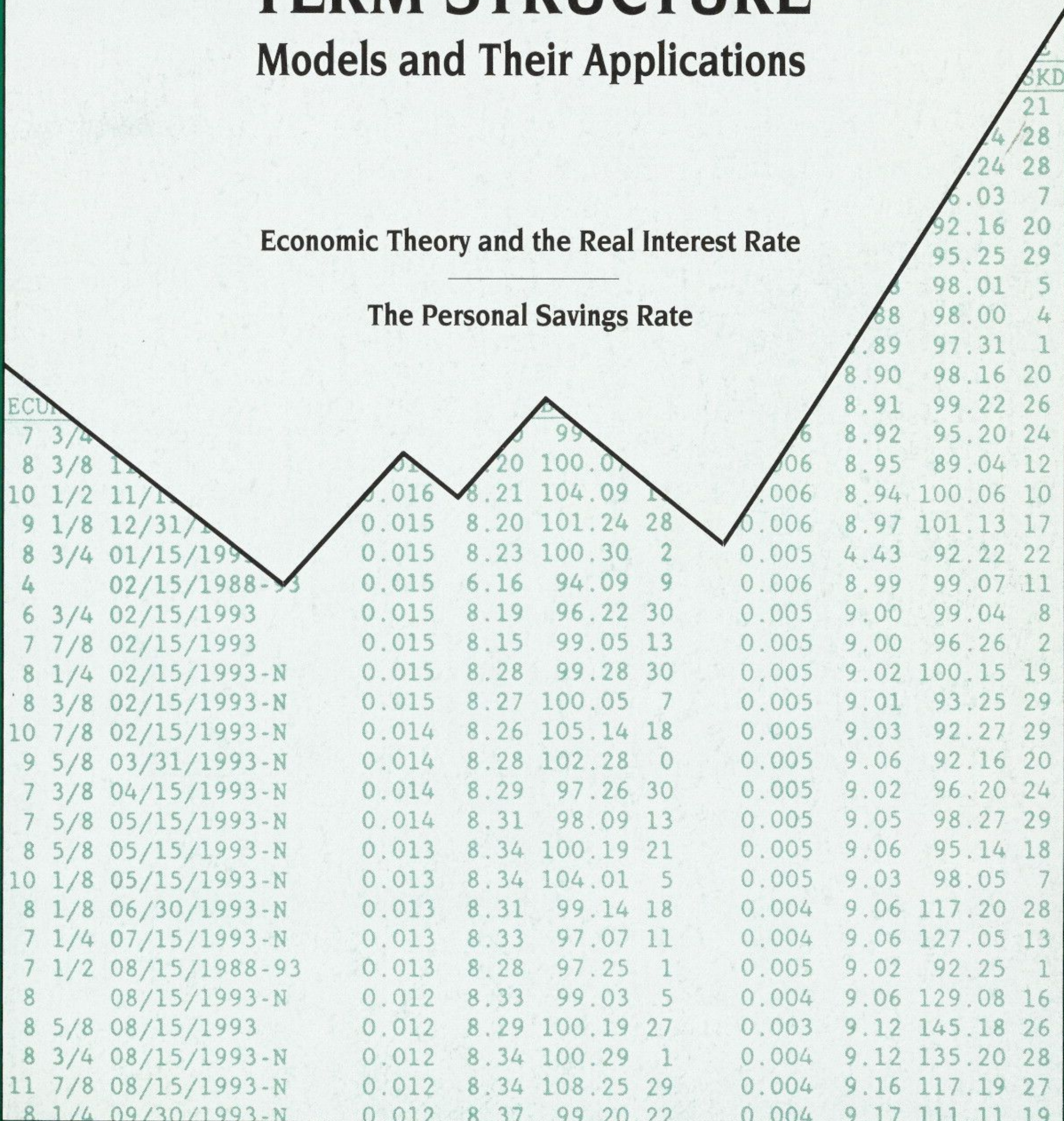
Economic Review

FEDERAL RESERVE BANK OF ATLANTA JULY/AUGUST 1990

INTEREST-RATE TERM STRUCTURE

Models and Their Applications



Economic Theory and the Real Interest Rate

The Personal Savings Rate

Economic Review

The **Economic Review** seeks to inform the public about Federal Reserve policies and the economic environment and, in particular, to narrow the gap between specialists and concerned laypersons.

Views expressed in the **Economic Review** are not necessarily those of this Bank or of the Federal Reserve System.

Material may be reprinted or abstracted if the **Review** and author are credited. Please provide the Bank's Public Information Department with a copy of any publication containing reprinted material.

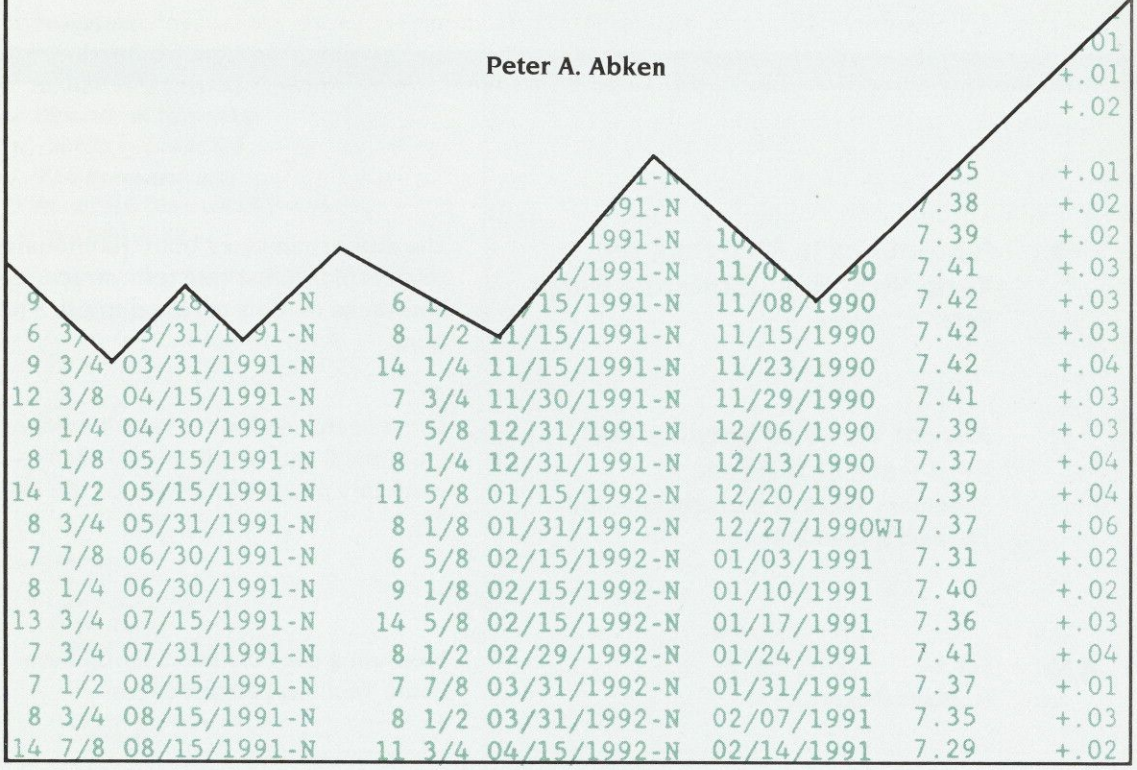
Free subscriptions and limited additional copies are available from the Public Information Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713 (404/521-8788).

Change-of-address notices and subscription cancellations should be sent directly to the Public Information Department. Please include the current mailing label as well as any new information.

ISSN 0732-1813

Innovations in Modeling the Term Structure of Interest Rates

Peter A. Abken



Economists and market practitioners have long sought to understand interest-rate movements. As a result, a number of term-structure models have been developed over the years. Major advances in option-pricing theory in the 1970s led the way to significant progress. In this article, the author discusses two recent term-structure models that have been highly influential in stimulating further research and are being applied to valuing a wide variety of financial instruments.

The phrase *term structure of interest rates* refers to the relationship between interest rates on bonds of different maturities. It is no doubt familiar to most who peruse the newspaper's financial pages. A precise understanding of what determines the relationships among these interest rates is still lacking, however, despite a voluminous amount of research. The reason for this open-endedness is not hard to grasp: current interest rates to a large extent reflect expectations of future interest rates, as well as all relevant factors that impinge on them.

Financial economists have long sought to characterize, and more importantly to predict, interest-rate movements using mathematical models. These models tend to shape thinking about the term structure even if a formal model is not a conscious part of this process. Policymakers, money managers, and other investors often look to the term structure for clues about the market's expectation regarding future interest rates. To make sense of the array of interest rates determined in financial markets, any such divination usually implicitly or explicitly uses a theory of the term structure. The interest-rate forecasts embedded in the term structure may help to form individual forecasts, upon which all kinds of economic decisions are based. Differences between individual and market outlooks on interest rates may also spur investors and speculators to bet against the market through their trading of bonds. This trading moves bond prices and thus leads to a melding of private information with the public information embodied in the term structure.

Since about the mid-1980s the most important application of term-structure models has been in valuing various kinds of interest-rate options. The most basic of these are bond options, which can be used to hedge the value of bond portfolios against capital losses. Interest-rate caps and swaps are examples of more complex interest-rate contingent claims that can hedge interest payments on debt whose interest rate fluctuates periodically with market rates.[1] Caps and swaps can be viewed as portfolios of bond options and may be valued using term-structure models. Much current financial research centers on refining existing term-structure models and on developing new ones that are easier to use and more accurate in their predictions.

This article begins with a review of the elementary theories of the term structure. These theories—the expectations hypothesis, the liquidity premium hypothesis, and the preferred habitat hypothesis—have been standards in economics since the 1960s and still constitute the core of contemporary textbooks. These theories are no longer "state of the art," however. This article attempts to bridge the gap between the traditional hypotheses of the term structure and more recent, less accessible work stimulated by innovations in options pricing theory since the early 1970s. Additionally, the discussion of the term-structure models explores the connection between the new modeling "technology" that produced the path-breaking Black-Scholes option-pricing formula published in 1973 and its applicability to pricing bonds, which, like options, are another kind of so-called contingent claim.

Intended as a nonmathematical exposition of both the traditional and recent models of the term structure, this article introduces the term structure by briefly reviewing the three traditional hypotheses as well as the newer models, which are essentially elaborations of the same concepts. Two recent-vintage term-structure models, developed in the mid-1980s, are examined in detail—the Cox-Ingersoll-Ross (1985b) and Ho and Lee (1986) models, which have been highly influential and represent different directions that modeling efforts have taken. To provide a context for these models, related models are also discussed. A brief survey of very recent research shows how these newer models have been extended or applied, and an option pricing example using the Ho and Lee model illustrates an important application of these models.

Whether pricing bonds or contingent claims, all models considered in this article share certain basic principles. The new models explicitly build in uncertainty about the course of future interest rates; they are models of random interest rates. Knowledge of constraints on the behavior of interest-rate movements allows for construction of models that value bonds or interest-rate contingent claims. This article explains how this valuation is accomplished.

The Term Structure under Uncertainty

The term structure of interest rates refers to yields on bonds that are alike in all respects except their time to maturity. Default

---

The author is an economist in the financial section of the Atlanta Fed's research department.

risk, for example, should be the same across all bonds to allow meaningful comparisons of bond yields spanning the maturity spectrum. Not only because credit risk is absent but also because government bonds offer the broadest range of maturities of all bonds, most analyses of the term structure are conducted using default-free government bonds. The box on page 5 gives some elementary definitions concerning zero-coupon bonds as well as the arithmetic for relating zero-coupon bond prices and yields to maturity. The box on page 12 considers the basic building blocks for the expectations hypothesis and explains how arbitrage forces equality in the holding-period returns on bonds of different maturities. The analysis shown has been conducted in a theoretical world of perfect certainty about future interest rates.

Expectations Hypothesis. Once the future becomes uncertain, investors face interest-rate risk.[2] This risk is irrelevant for investors whose future cash needs exactly match their bonds' maturity dates. The bonds mature at par, and no capital gains or losses are possible. Such a coincidence is the exception rather than the rule, though. If investors need cash sooner, they are necessarily exposed to the possibility of capital loss; that is, newly issued bonds for the same maturity date may bear higher yields, forcing existing bond holders to sell at a discount to match the higher yield. Conversely, if their bonds mature sooner than the time of their cash need, the investors risk reinvesting at a lower yield in a new bond that matures on the desired date. If they had bought a longer-term bond at the outset, no such risk would have been incurred.

In economics the existence of risk implies that the outcome of a decision is not known precisely; however, a range of possible outcomes is known, each of which has some chance—some probability—of occurring. Thus, investors are assumed to know the range of outcomes and, more specifically, the probability distribution for the likelihood that any particular outcome will occur. These and other statistical concepts are fundamental to theories of the term structure because they quantify investors' expectations.

In the expectations hypothesis of the term structure, risk has no effect on investors' choices. All that matters is their expectation of future interest rates at all maturities. Here the term *expectation* is used in its narrow statistical sense to mean average or mean value for a particular future interest rate. The range of outcomes for an interest rate may be quite wide or quite narrow, but that issue is irrelevant to investors. Investors are said to be "risk neutral."

The expectations theory posits that, regardless of the holding period considered, any possible combination of bonds must offer the same rate of return. This idea, illustrated in the box on page 12 for one- and two-year bonds, works exactly the same way for any other maturity combinations. Any long-term bond yield is the average of the current short-term yield and future expected short-term yields. The essential assumption is that investors care only about holding-period returns and consider all bonds of different maturities as perfect substitutes. All interest rates along the yield curve are therefore linked by arbitrage.

Liquidity Premium Hypothesis. Rather than "risk neutral," investors may be "risk averse," in which case their average sensitivity to bearing risk will be reflected in bond prices. The usual presumption is that longer-term bonds are riskier in the sense that their principal is more exposed to interest-rate fluctuations over time. Specifically, this exposure makes them less liquid (less readily converted to cash by selling them) because they bear a greater potential for capital loss before maturity. Risk-averse investors therefore are believed to require a "liquidity premium" to induce them to hold longer-term bonds willingly. The liquidity premium manifests itself as an increase in the forward rate above the future spot rate, which results in a shifting of the yield curve above the one predicted by the expectations hypothesis. Known as the liquidity premium hypothesis, this concept is a refinement of the expectations hypothesis. The liquidity premium is alternatively known as a *risk premium*, which will be the term used hereafter. To distinguish it, the expectations hypothesis is sometimes called the pure (or risk-neutral) expectations hypothesis.

The liquidity premium hypothesis is rather vague about what determines the risk premi-

Basic Terminology

The basic financial instrument of interest is a zero-coupon bond that promises a fixed payment, which will be assumed to be $1, at a future date.[1] The bond makes no interim payments, that is, it has no coupons. While zeros—Treasury bills and zero-coupon Treasury bonds—trade in the bond markets, many bonds make periodic coupon payments. Because any coupon bond can be decomposed into a zero coupon bond by treating each future coupon payment as well as the repayment of principal as a separate zero coupon, modeling in terms of zeros is not unrealistic.[2] The models to be considered also abstract from credit (or default) risk so that the uncertainty stems only from unknown future interest rates.

A zero-coupon bond is alternatively known as a discount bond because its price is less than its $1 face or par value. Its return is derived strictly from the appreciation in price at the time of maturity. For the sake of simplicity, bonds are assumed to be issued at yearly intervals with maturities in multiples of a year, making the shortest maturity bond a one-year bond. The current date will be time 0.

The bond price, B, is the price at time t, when the bond is purchased, for principal repayment upon maturity at later date T. The difference between t and T is called the bond's time to maturity. Consider a bond with one year to maturity (a one-year bond) issued today at time 0. The bond's yield to maturity, $R(1)$, is determined from the following equation involving the bond's price:

$$1/B(1) = 1 + R(1), \text{ or}$$
$$B(1) = 1/[1 + R(1)].$$

The time to maturity is enclosed in parentheses. If the time to maturity is one year and the yield $R(1)$ is 8 percent, then $B(1)$ is approximately 0.926. The term $R(1)$ is expressed as a simple interest rate. Clearly, yield and bond price are inversely related: as one goes up the other goes down.[3]

Consideration of multiple periods entails compounding of interest. The yield to maturity, $R(n)$, of an n period zero-coupon bond is

$$1/B(n) = [1 + R(n)]^n, \text{ or}$$
$$B(n) = 1/[1 + R(n)]^n.$$

The term $R(n)$ is also expressed as a simple annual interest rate. Plotting the yield to maturity against time to maturity gives the so-called yield curve. The reasons for its shape or structure are the subject of the various hypotheses discussed in the main text.

Notes

[1] Of course, any par value can be used simply by multiplying the bond price by the face amount. For example, a bond promising $10,000 upon maturity would be worth 10,000 times the purchase price of the $1 par value discount bond.

[2] See Bodie, Kane, and Marcus (1989, 420) for a procedure used to infer zero-coupon bond prices from the prices, coupons, and principal payments of coupon bonds.

[3] For a fuller discussion of this terminology as well as the basics concerning the traditional hypotheses of the term structure, the reader is referred to Bodie, Kane, and Marcus (1989). Almost any undergraduate investments or money-and-banking textbook contains some discussion of these topics.

um, other than supposing that such a premium is the investor's reward for bearing interest-rate risk. The new models incorporating this hypothesis give more structure to the underlying determinants of the risk premium.

Once risk aversion is assumed to affect interest rates, interpretation of the term structure becomes problematic. Without knowing the size of the risk premium, the expectations hypothesis cannot be used to infer future short-term interest rates. Any forward rate now consists of a risk premium and the expected short-term interest rate. Consider the simple example of deciding between an investment in a sequence of consecutive one-year bonds or an investment in a two-year bond, like that used in the box on page 12. Assume that the one-year yield is 5 percent and the future one-year yield is known to be 7 percent. If the true risk premium were known to be 0.5 percent (or 50 basis points), the two-year yield would be:

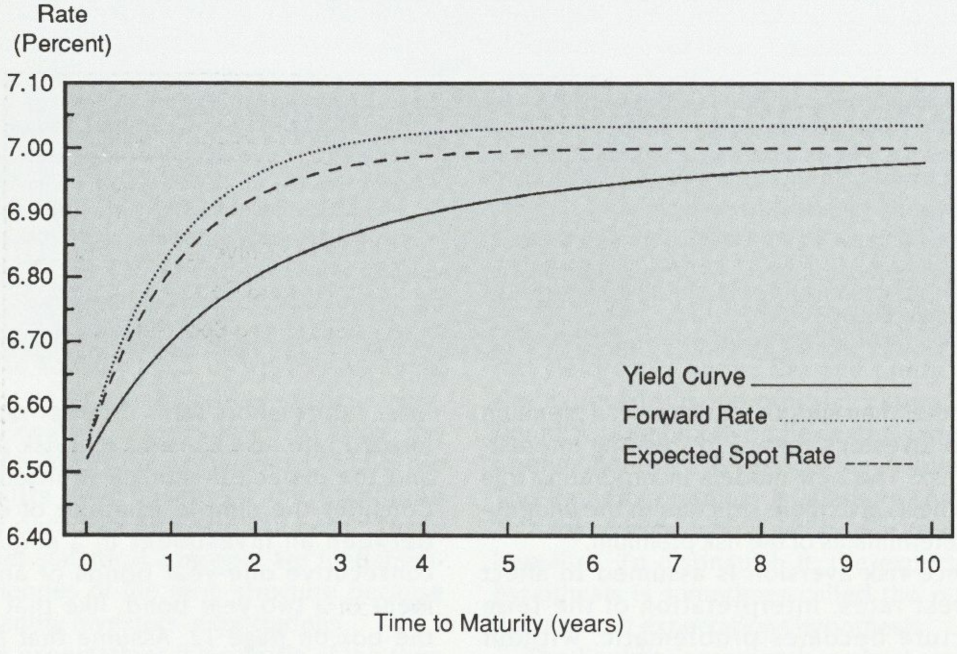
$$[5\% + (7\% + 0.5\%)]/2 = 6.25\%,$$

that is, the average of the short rate and the forward rate (which is the sum of the future short rate and the risk premium). The two-year yield is 6.25 percent instead of 6 percent as in the risk-neutral case (where the risk premium is zero). If the risk premium is unknown, all that can be inferred from the current term structure (one- and two-year bond yields) is that the implied forward rate is 7.5 percent. More information is required to disentangle the risk premium from the expected short rate.

For a hypothetical continuum of interest rates along the yield curve, Chart 1 shows the relationship among yield curve, expected short-term interest rates, and forward rates. If the term structure slopes downward—that is, short-term bond yields are above long-term bond yields—one may safely infer that expected short-term rates are lower than the current rate, even if there are risk premia in the forward rates. However, upward-sloping yield curves do not imply that expected short-term rates are rising, since flat or declining expected short-term rates may be sufficiently augmented by rising risk premia to produce rising forward rates. Chart 2 illustrates this phenomenon.[3] These interpretations will be discussed again in the review of recent theoretical models of the term structure.

Preferred Habitat Hypothesis. A third traditional theory is the preferred habitat hypothesis, which posits that interest rates along the yield curve result from market forces in different maturity segments. Sufficiently large risk premia or discrepancies between investors' and market expectations of interest rates may lure investors away from their preferred habitat. Thus, as Franco Modigliani and Richard Sutch (1966) have observed, "risk aversion should not lead investors to prefer to stay short but, instead, should lead them to hedge by staying in their

Chart 1.
Liquidity Premium Theory



maturity habitat, unless other maturities (longer or shorter) offer an expected premium sufficient to compensate for the risk and cost of moving out of one's habitat" (184).[4]

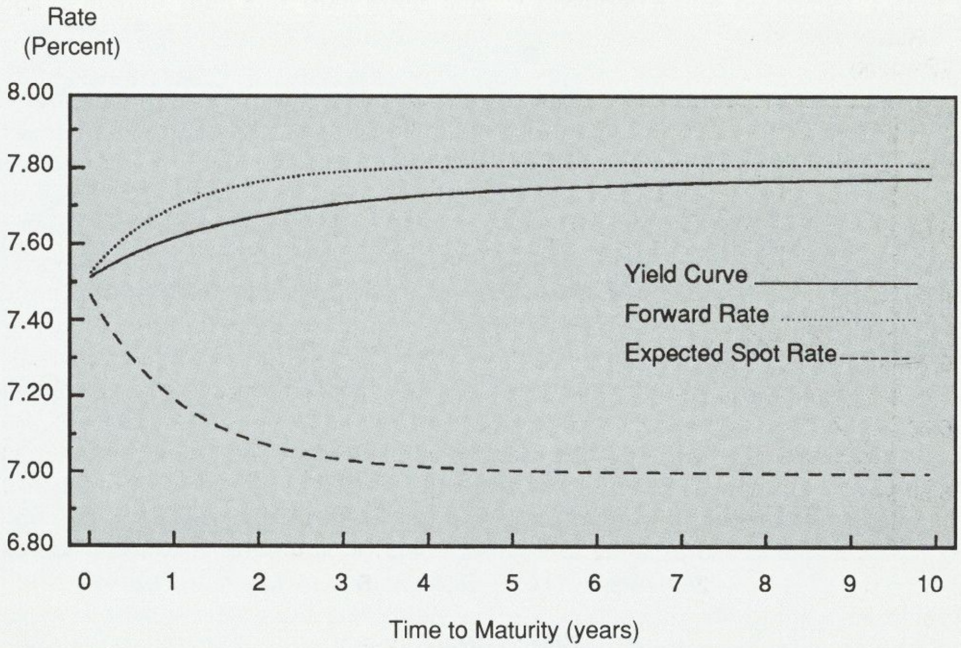
Life insurance companies' and pension funds' typical preference for investments in longer-term bonds exemplifies preferred habitats, as does depository institutions' likely penchant for short-term bonds. Life insurance companies and pension funds have relatively predictable long-term liabilities, which they match against investments in long-term bonds and other long-term assets. Similarly, depository institutions tend to fund relatively short-term loans with short-term liabilities. Their bond holdings, therefore, also tend to have short-term maturities.

The supply and demand for bonds within each habitat is assumed to have an impact on interest rates; bonds of different maturities are not considered perfect substitutes. In short, institutional characteristics, not just interest-rate expectations, play a role in determining the term structure. An earlier version of this theory, the market segmentation hypothesis, assumed a rigid segmentation of markets, which is now generally regarded as implausible. The preferred habitat hypothesis synthesizes the expectations and market segmentation hypotheses.

Innovations in Modeling the Term Structure

In 1973 Fischer Black and Myron Scholes published their pathbreaking article, "The Pricing of Options and Corporate Liabilities," which transformed not only academic financial research but also actual financial practice. Their celebrated formula allowed academics and practitioners alike to price all kinds of contingent claims.[5] The theory proved to have very broad applications. John C. Cox, Stephen A. Ross, and Mark Rubinstein (1979) have noted that option-pricing theory "applies to a very general class of economic problems—the valuation of contracts where the outcome to each party depends on a

Chart 2 . Liquidity Premium Theory



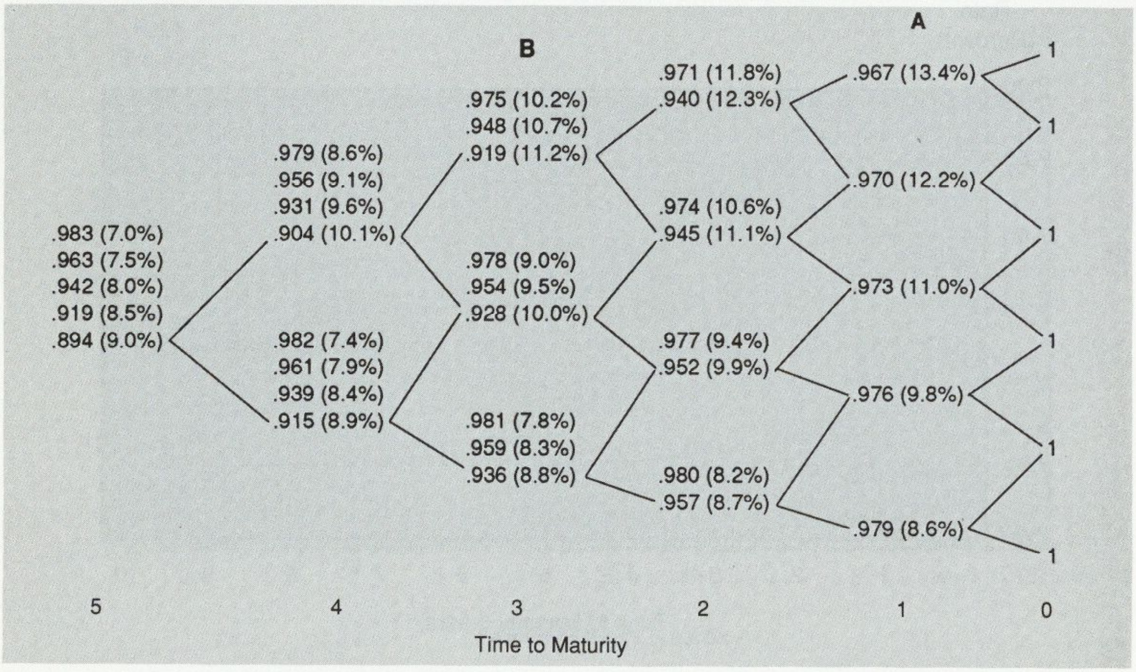
quantifiable uncertain future event" (230). Term structure modeling is one area of financial research that has benefited from the advent of modern option-pricing theory. A detailed review of two general approaches, using two specific models, will develop the connection between option-pricing theory and term-structure modeling. The two models highlight the basic directions of recent research.

The Ho and Lee Model. Proposed by Thomas S.Y. Ho and Sang-Bin Lee in 1986, the Ho and Lee model is a useful starting point in the exposition of term-structure models because it uses what is called a lattice approach. The lattice approach to option pricing explicitly conveys what is meant by a "quantifiable uncertain future event." The basic problem to model is how the price of a zero-coupon bond changes from the present to its future maturity date. Only at the present moment and at the future maturity date is the bond price known with certainty. The Ho and Lee model is one attempt to quantify the evolution of bond prices over time. As will be illustrated below, this quantification is essential to pricing contingent claims on bonds.

Ho and Lee abstract from the complexities of actual bond price movements by assuming (1) that a bond price moves only at fixed intervals over time (for example, every day or every minute) and (2) that they move either up or down when they change. The up or down price changes have an associated probability of occurring, of, say, 20 percent and 80 percent, respectively. Thus, starting from some known price at the current date, future price changes are restricted to evolve by successive up or down movements, which trace out a tree or lattice pattern.

Chart 3 illustrates the branching process that begins with the initial discount bond price. The initial bond in the example has fifteen months (five quarters) to go before maturity, when it will pay $1. The yield to maturity is assumed to be 9 percent, and the current price is $0.894. The annualized yield to maturity appears in parentheses. Succes-

Chart 3.
Evolution of Bond Prices in the Ho and Lee Model
(yields in parentheses)



Note: The one-period bond appears at the top of each column; time to maturity increases by one period with each successive row.

sively shorter-term bond prices and yields appear above the 9 percent, fifteen-month bond. Thus, the entire term structure is given in the far left-hand column (five quarters to maturity), starting at the top of the list with a 7 percent, three-month bond that matures (and disappears) in the next column and extending to the longest maturity bond, the 9 percent bond at the bottom of the list. The illustration is therefore given for an upward-sloping term structure.

The tree fans out from the initial price as up and down possibilities proliferate at each future point in time. Each path through the branches to any particular vertex or node in the tree diagram represents a particular succession of up and down price movements. These movements correspond to the realizations of particular "states," which are the possible random outcomes schematically depicted in the tree diagram. After one period elapses from the initial date, two states are possible; after two periods, three states; and so forth. In other words, time is measured horizontally, although states are indicated vertically in the tree diagram. The tree therefore depicts the value of a particular bond in all states as it approaches maturity. In this example, the bond initially has five periods to mature, then four, then three, and so forth. The final branches in the diagram end with $1 in all states, since the bond's value at maturity is known with certainty. The branches link the path of the initial five-quarter bond as it approaches maturity. The shorter-term bonds could also be joined in this fashion to show their progression.

The essential idea behind the evolution of the bond price tree is that the forward rates implied by the initial term structure would actually be the future short-term rates to prevail in the absence of disturbances or shocks causing changes in interest rates. In the Ho and Lee model, bond prices throughout the tree result from perturbations of the initial implied forward rates. Any state price at a given time represents the initial implied forward rate altered by an accumulation of up and down shocks, resulting in a particular position in the tree diagram. Moreover, the size of the price changes, governed by the perturbation function, is restricted in such a way that no arbitrage profits can be realized; that is, the internal consistency of the model requires that no arbitrary portfolio of discount bonds of different maturities can be formed that earns more than the risk-free rate when the portfolio is perfectly hedged (risk free).

To see how prices are related in the tree diagram, consider any two adjacent state prices. The node marked "A" in Chart 3 is derived as the discounted value of the two certain $1 payoffs at the maturity date, each of which has an equal chance of occurring in this example (that is, the probability is 0.5). The discount factor of 0.967 at node A is in fact the price of a one-period discount bond. This discount factor was computed using Ho and Lee's formula for the perturbation function evaluated at this particular node.[6] All state prices for time 1 are one-period prices. For all earlier periods the bond prices are computed using the same recursive procedure. At node B the price is again the weighted average of two future bond prices times the one-period bond price:

$$(0.5 \times 0.940 + 0.5 \times 0.945) \times 0.975 = 0.919 .$$

This calculation represents a discounted expected value since at node B there are two possible future outcomes. Either the bond price rises to 0.945 or falls to 0.940. The average of these is discounted to time 3 using the one-period bond price for time 3. Again, the discount factor is the one-period bond price, 0.975, determined one period earlier (at the top of the bond price list for node B). In summary, the tree diagram represents an expected value calculation, for which the known initial bond price of 0.894 is the final outcome.

At first glance the tree diagram may not seem very useful since the final computation is a bond price that was already known at the outset. However, the important aspect of this exercise is that the various price paths (possible branching patterns) are fully described in the tree. This quantification of a bond's future state prices is essential for contingent claims pricing.

A characteristic of this type of lattice model is that the order of up and down movements does not matter; instead, the cumulative number of up (or down) moves from the initial node determines the price at any future time-state node. Inspection of the tree reveals that, from any interior node, moving rightward first up and then down is identical to moving down, then up. This restriction plays an important role in valuing bonds in this model. Also, the initial upward-sloping term structure retains its slope regardless of its location in future periods, but the levels change. In fact, all interest-rate movements are perfectly correlated in the Ho and Lee model. More complex models that will be considered below avoid this unrealistic feature.

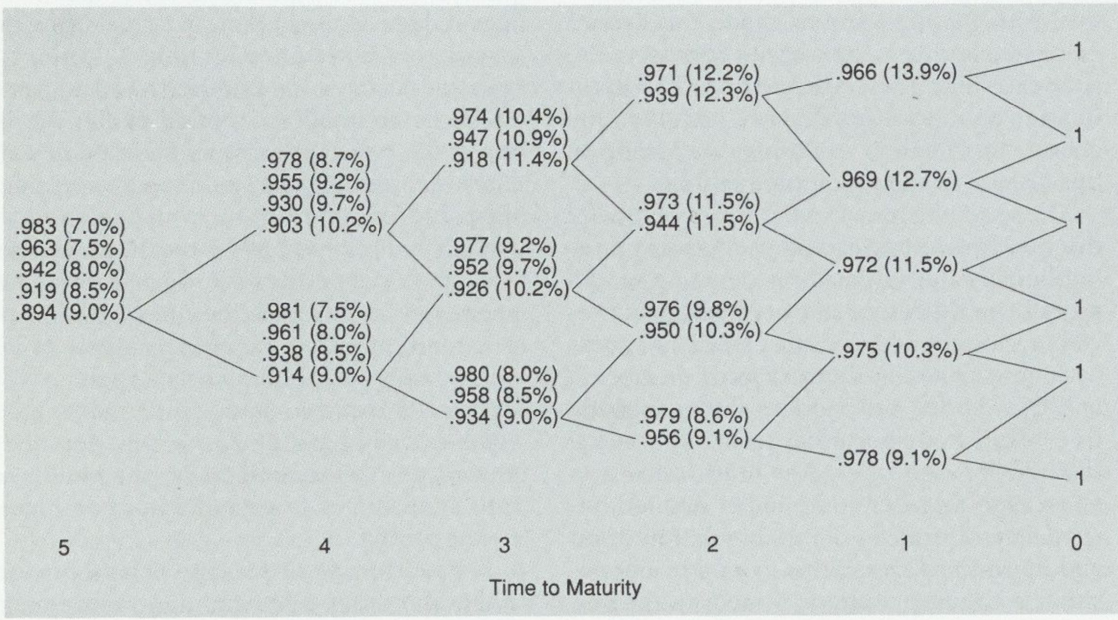
Another drawback of the model is the fact that, depending on the model's parameter values, some bond prices may exceed face value before maturity. This outcome along a price path implies that the bond's interest rate is negative. Peter Ritchken and Kiekie Boenawan (1990) prove that such an occurrence represents an internal inconsistency of the model and show how to modify the model to preclude negative interest rates. Charts 3 and 4 were produced using the Ritchken-Boenawan modification.

Despite its limitations, the Ho and Lee model has the virtue of clearly and simply illustrating how the current term structure is derived as the discounted expectation of the future bond payoffs. More recent models of this type rectify some of these limitations and consequently are transforming this modeling approach into a more useful tool for applications as well as theory.

A Shift in Risk Aversion. The term-structure movements depicted in Chart 3 are assumed to be generated in a world in which investors do not need to be compensated for bearing interest-rate risk. If investors' preferences shifted to become risk averse with respect to interest-rate risk, what effect would that have on the evolution of bond prices? In general, the expected rate of return for an asset would have to increase for the investor to continue holding that asset willingly. For illustrative purposes, the probability of an upward move was assumed to be the same as the probabil-

Chart 4.
Ho and Lee Model Shift in Risk Aversion Bond Prices

(yields in parentheses)



Note: The one-period bond appears at the top of each column; time to maturity increases by one period with each successive row.

ity of a downward move in Chart 3. The increase in risk aversion is quantified as an increase in the probability of an upward move from 50 percent to 60 percent. These probabilities are arbitrarily chosen to show the change in the evolution of bond prices. Any increase in the up-move probability would accomplish the same end.[7]

The Ho and Lee model takes the initial term structure as given in the course of valuing future bond prices and contingent claims. Consequently, a shift in preferences is considered to occur after the initial period. Chart 4 gives the new tree diagram, starting with the same initial term structure as Chart 3 and also ending with bond prices equal to par upon maturity in period 5. However, the rate of price appreciation in Chart 4 is always greater (or bond prices are everywhere lower), regardless of the path taken through the tree diagram. The terminal payoff of $1 is the same in both diagrams. The downward shift in prices across all states and times occurs because the discount rates have increased because of the rise in risk aversion. In other words, interest rates are always greater in Chart 4, which reflects the risk premium now included in rates.

The Cox-Ingersoll-Ross Model. The Cox-Ingersoll-Ross model takes an approach to valuing fixed-income securities and their contingent claims that is fundamentally different from Ho and Lee's. The basic difference is that the former is a general equilibrium model, whereas the latter is a partial equilibrium model. Essentially, the Cox-Ingersoll-Ross approach has a deeper theoretical foundation and hence more ambitious goals in terms of economic modeling. The practical distinction is not as sharp at the level of using these models to value bond options.

The Ho and Lee model is a partial equilibrium model in the sense that the initial term structure and the process generating shifts in that structure are assumed to be given outside of the model—that is, they are determined *exogenously*. On the other hand, as will be discussed below, the Cox-Ingersoll-Ross model takes investor preferences and unforecastable shocks to physical investment opportunities as given. That is, term-structure movements themselves are explained within the context of the model; they are explained *endogenously*. The importance of this general equilibrium formulation will become apparent after some preliminary discussion.

The Cox-Ingersoll-Ross model can be stripped of its underlying general equilibrium structure, rendering it a partial equilibrium model. This simplification is a useful starting point for understanding this type of model. In fact, a number of similar models originated at about the same time as the Cox-Ingersoll-Ross model, but these were conceived as partial equilibrium models.[8]

Another difference between the Cox-Ingersoll-Ross and Ho and Lee models is that the former uses continuous-time mathematics, which demands considerably greater technical sophistication than does the discrete-time mathematics of the Ho and Lee model. The use of continuous-time mathematics requires assumptions that all bonds and other financial instruments trade at every moment in time—that is, continuously—and, furthermore, that bonds mature at every moment in time. While highly artificial, this kind of model allows the tools of continuous-time mathematics to deliver pricing formulas for bonds and other financial instruments. In fact, the original Black-Scholes formula was derived using continuous-time mathematics.[9]

The simplest version of the Cox-Ingersoll-Ross model assumes that the term structure can be expressed as a function of one variable, the instantaneous risk-free rate or spot interest rate, that is, the interest rate on a bond that matures at the next "instant" in time. The continuum of bond prices along the term structure is expressed solely as a function of the spot rate and time (or, equivalently, time to maturity). All shocks to the term structure therefore emanate from the spot interest rate, which in fact is a proxy variable for the fundamental uncertainty in the economy. To give more content to this formulation requires appending the full general equilibrium structure of the economy to the model. In contrast, the partial equilibrium model simply posits the instantaneous rate as the source of random fluctuations in the term structure. The spot interest rate is referred to as a "state variable," a variable that summarizes information about uncertainty in the

The Term Structure under Certainty

The question to be answered is, how is the yield on a one-year bond (the "short" rate) related to the yield on a two-year bond (the "long" rate)? Alternatively phrased, for a two-year investment horizon, what governs the decision to select a sequence of two consecutive one-year bonds or one two-year bond? Of course, the answer is that one chooses the alternative that offers the highest yield, but this response must include an important qualification. The choice hinges on what the investor expects the yield will be on the one-year bond to be issued one year hence. Given that expectation, the investor would choose the higher-yielding alternative. However, if there were uncertainty about the future one-year yield, the investor might choose the certain yield on the two-year bond instead. For now we assume that investors hold their expectations with certainty. Certainty means that the expectation will actually be realized in the future.

Investors obviously have strong incentive to seek out the highest return on investments of a given riskiness.[1] This behavior influences the term structure by bidding bond prices, and therefore yields, to a level that reflects the market sentiment about future interest rates. For example, suppose that an investor observes the one-year bond to be currently yielding 5 percent and the two-year bond, 6 percent annually and that the investor believes that the future one-year yield will be 8 percent. To make an investment decision, the investor must determine what the future short rate is implied by market interest rates. For the investor to be indifferent about the investment alternatives over the two-year horizon, the total holding-period rates of return must be equal, that is,

$$2 \times 6\% = 5\% + \text{future short rate};$$
$$\text{future short rate} = 7\%.$$ [2]

In other words, the fixed annual yield on the two-year bond summed over two years must equal the sum of the short rate and the expected future short rate. Although the future short rate is not directly observed, it is implied by the one-year and two-year bond rates and is usually referred to as the implied forward rate. In this example assuming certainty about expectations, the forward rate is the same as the expected future short rate.

Given his or her expectation that the future short rate is 8 percent, the investor would rationally choose to buy the current one-year bond and roll over that investment into the future one-year bond. This sequence of bonds would be equivalent to a two-year bond currently offering about 6.5 percent, as compared with the actual yield of 6 percent. The reason that the two-year bond is currently mispriced at 6 percent could be, for example, that investors received news leading them to expect higher inflation a year from now (raising their expected short rate). Their investment choice bids down the price of the two-year bond and hence bids up its yield. Once that yield reaches 6.5 percent, the bond market is back in equilibrium—investors are again indifferent about the investment alternatives. On the other hand, if the expectation were, say, for a 5 percent future one-year bond yield, investors would choose the two-year bond and consequently bid down its yield.

In this example, investors are said to have exploited an arbitrage opportunity. They have profited from the temporarily mispriced two-year bond without taking any risk.[3] Admittedly, this example is highly artificial because expectations are held with certainty; hence, there is by definition no risk. However, this kind of analysis carries over into a more realistic world characterized by uncertainty.

Notes

[1] All bonds are simply assumed to exist and hence no consideration is given to borrowers' behavior, that is, to the decision to issue bonds of particular maturities. Since the discussion applies to default-free bonds—namely, Treasury securities—ignoring the issuer does not detract from the analysis.

[2] Allowing for compounding of interest, the calculation is as follows: $(1.06)^2 = (1.05)(1 + \text{future short rate})$. The future short rate is therefore approximately 7.0095 percent. Thus, the arithmetic average used in the text is fairly accurate in this example.

[3] For the above example, the arbitrage was indirect since it involved selecting the better investment. A standard arbitrage transaction for this case would entail buying the higher-yielding investment alternative by selling the lower-yielding one; that is, arbitrage profits would be earned with a zero net

investment. Profits are ensured since the yields must eventually equilibrate.

This version of the example assumes that borrowing and lending rates are equal and that the proceeds from selling the lower-yielding bond "short" are immediately and fully available to buy the higher-yielding bond. In general, arbitrage connotes *risklessly* profiting from buying a good at a lower price and selling it at a higher price in the same or another market.

economy. More complex models can be constructed by including more state variables, each of which convey independent bits of information about shocks to the term structure. A fuller interpretation of state variables will be possible once general equilibrium models are considered.

The Cox-Ingersoll-Ross model restricts the behavior of the spot rate by supposing that, although random in its movements from one instant in time to another, the spot rate tends to "revert" to a long-term level and never becomes negative. It is prevented from becoming negative because negative real interest rates are not economically meaningful in the context of their model. This restriction alone is not sufficient for a realistic model since the spot rate would exhibit too much variability over time. Thus the assumption of mean reversion was incorporated. Note that the original Ho and Lee model, cast in terms of nominal bond prices and interest rates, does allow negative interest rates and does not have a mean-reversion property.[10]

Valuing a bond using the Cox-Ingersoll-Ross model, like the Ho and Lee or any other valuation model, entails computing the expected value of the discounted payoff. Instead of positing a discrete number of states upon expiration of a bond as in Ho and Lee, the Cox-Ingersoll-Ross model allows for a continuum of outcomes. Again, bond valuation means determining the expected value of the discount factors to be applied in the continuum of states, since the maturity value of a bond is always par. The chief obstacle to making this determination is that the discount factors are unknown for risk-averse investors. Fortunately, the option-pricing methodology that originated with Black and Scholes provides an ingenious solution.

Risk-Neutral Valuation. With certain assumptions, option valuation can be formulated in such a way that knowledge of investors' risk aversion (that is, their discount functions) becomes irrelevant. This approach applies to many kinds of option-pricing problems though, unfortunately, not directly to term-structure applications for reasons to be discussed shortly. Although step-by-step details pertaining to option valuation are beyond the scope of this article, some background on the manner in which risk aversion is treated in valuation will aid the understanding of term-structure models. If certain assumptions are imposed—continuous trading, continuous price dynamics, and the property of nonsatiation (that is, that investors always prefer to consume more rather than less)—valuation may proceed as if in a risk-neutral world, one in which the discount rates contain no risk premia and thus are observable.[11] Valuation is said to be preference free. This line of reasoning has become known as the Cox-Ross risk-neutrality argument, first expounded in Cox and Ross (1976), and is a useful starting point for considering valuation of interest-rate contingent claims.

The validity of the risk-neutrality argument depends on the construction of a so-called hedge portfolio consisting of the derivative asset (for example, options) and underlying assets (like stocks) that are traded in such a way that the portfolio is without risk. The hedge portfolio would be constructed so that, for instance, a rise in the value of the stock component would be exactly offset by a fall in the value of the option component and vice versa. Because the portfolio is riskless, any funds invested in it would have to earn the risk-free rate of interest; otherwise, arbitrage would be induced between the hedge portfolio and the risk-free asset. The value of the portfolio's components can be determined without regard for risk premia. No matter what the actual degree of risk aversion, the fact that the portfolio is riskless means that the expected rate of return on its components can be assumed to be the riskless rate as well. In other words, the same option price would be derived for any discount rates (if they were known). Thus option pricing proceeds using the preference assumption that makes valuation easiest: risk neutrality.[12]

Allowing for random interest rates complicates the valuation process. The crucial aspect of the risk-neutrality argument is that the stock or underlying asset is itself the state variable and is a traded asset. However, the instantaneous interest rate is not a traded asset and therefore cannot be used directly in constructing a hedge portfolio. Unlike the hedge portfolio used in deriving the Black-Scholes option pricing equation, risk cannot be eliminated in such a way that valuation is preference free.

Nevertheless, the problem of valuing bonds of differing maturity, each of which may require a different and unobservable expected rate of price appreciation (in order to be held willingly by the investor), can be simplified. It is usually assumed that the bond price depends only on the state variables (in this case, the instantaneous interest rate) and time. A hedge portfolio can be formed, but doing so requires knowing something about risk preferences in the economy. In particular, for the Cox-Ingersoll-Ross model forming a riskless hedge portfolio implies that all bonds have the same return-to-risk ratio in equilibrium. That ratio reflects risk preferences. The basic idea is that any two bonds of arbitrary maturity can be traded in such a way (in a hedge portfolio) that they are a perfect substitute for any other bond of arbitrary maturity.

The excess return for a bond of one maturity (over the risk-free rate) relative to that bond's volatility (the standard deviation of its rate of return) must in equilibrium equal the excess return on another bond of different maturity relative to its volatility. This may be expressed mathematically as:

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2} = \lambda(r,t),$$

where μ_1 and μ_2 are the expected bond returns, σ_1 and σ_2 are the bond volatilities, r is the risk-free rate, and λ is the market price of instantaneous interest-rate risk.[13] The function λ can be a function of the state variable and time (though often it is assumed to be a constant), but it cannot be a function of any bond's maturity since all bonds are related to λ by the above equation. If this relationship did not hold for bonds of any maturity, arbitrage would be possible. Investors would always choose the bond with the highest excess return-to-risk ratio, thereby raising the bond's price and reducing its excess return until the bond's excess return-to-risk ratio equals λ.

Hence, the simplification used in the Cox-Ingersoll-Ross model and other continuous-time term-structure models is that, rather than needing to know the exogenous expected return for each bond, the only exogenous element that needs to be identified is the market price of interest-rate risk, which by hypothesis is shared in common by all bonds.[14] This parameter must be estimated from actual bond price data to use the Cox-Ingersoll-Ross or other similar model. Once the market price of interest-rate risk is estimated, valuation proceeds using risk-neutral pricing methods. The end result is a formula to price bonds of any maturity. In the context of the Cox-Ingersoll-Ross model, the bond price is solely a function of the spot rate of interest and the bond's time to maturity. In more general versions of the model, the bond price is a function of the underlying state variables and time to maturity. To summarize, interest-rate contingent claim pricing does involve making a risk adjustment in arriving at a pricing formula, but valuation is not preference free.

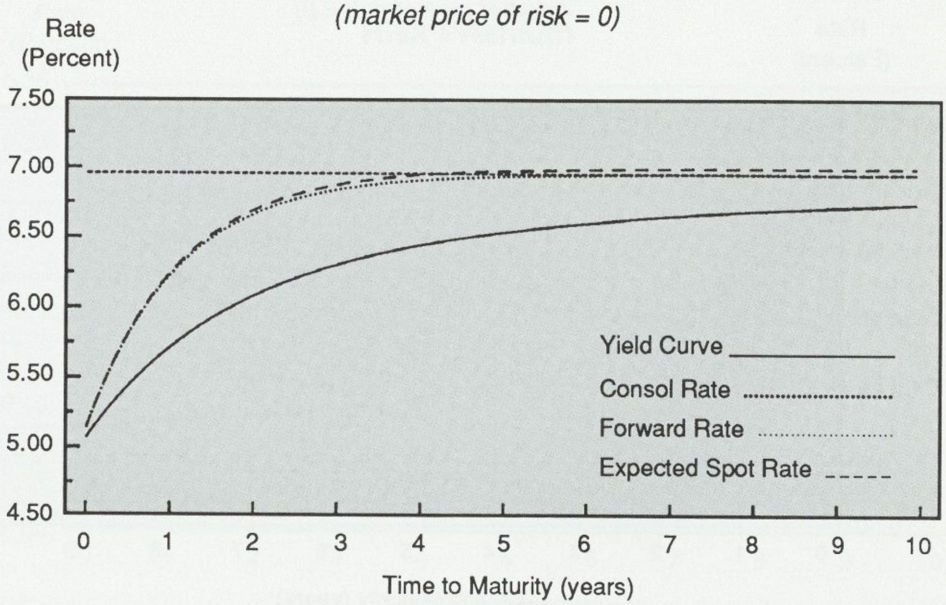
Speaking in terms of risk-neutral valuation in the context of term-structure models is something of a misnomer. Investors are not necessarily risk neutral when λ is zero. They may be risk averse, but, because of the uncertainty resulting from randomness in the spot interest rate (or, more generally, uncertainty stemming from the state variables), their risk aversion is such that they do not require a risk premium. These risk-averse investors' "portfolio decisions are completely myopic being made with no regard for hedging against changes in the state variables [which affect future output and consumption]" (Cox, Ingersoll, and Ross 1981, 783). Consequently, all future payoffs are discounted using risk-free rates, just as in those cases where the Cox-Ross risk neutrality argument applies. In pointing out this subtlety, Cox, Ingersoll, and Ross (1981) chose to call the case in which the factor risk premia are zero the "local expectations hypothesis" rather than the risk-neutral expectations hypothesis.[15]

An Illustration of the Cox-Ingersoll-Ross Model

Charts 5 and 6 are examples of the single-factor Cox-Ingersoll-Ross model, the simplest of their models. Its key features are illustrated in the graphs. The spot interest rate reverts to a long-term rate of 7 percent. Since expected spot rates are above the observed yield curve, any initial spot rate below the long-term rate results in an upward-sloping term structure. Any yield along the term structure can be expressed as a continuous-time average of the current spot rate and the expected spot rates. This relationship has a precise formula in the Cox-Ingersoll-Ross model. The expected spot rates were computed by formula in graphing their curves in Charts 5 and 6. Chart 6 shows a spot rate that is above its long-term rate, and consequently the term structure slopes downward. The rate at which the spot rate reverts to its long-term level is an important parameter that governs how the term structure moves over time. The other critical parameter affecting the term structure is the volatility of the spot rate.

Another feature of the single-factor Cox-Ingersoll-Ross model is that bonds of sufficiently long maturity have yields that are independent of the spot rate. This long-term yield, denoted as the consol rate (the yield on an infinite-maturity bond), appears as a heavy dotted line in all of the relevant charts. As maturity increases, the yield curve approaches this limiting yield "asymptotically," that is, very gradually. In contrast, multifactor models do not have the unrealistic characteristic that long-term bond yields are constant. The model developed by Michael J. Brennan and Eduardo S. Schwartz (1979) and another by Stephen M. Schaefer and Schwartz (1984) are two-factor models that include the consol

Chart 5.
Cox-Ingersoll-Ross Model
(market price of risk = 0)



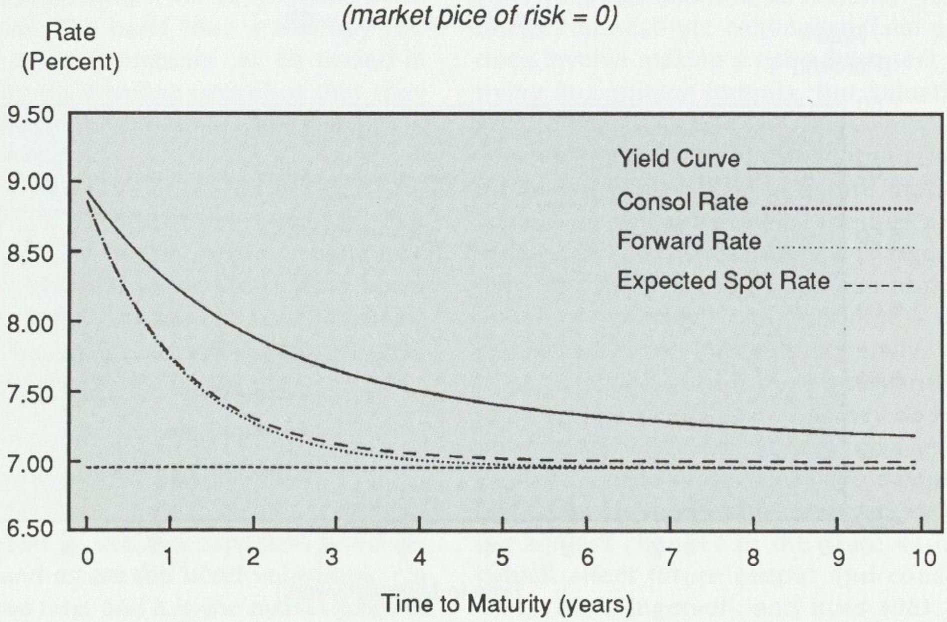
rate as an instrumental variable. This kind of model allows randomness to affect the term structure from opposite ends of the maturity spectrum. In fact, Ren-Raw Chen and Louis Scott (1990) have found that allowing for two and three factors greatly improves the fit between term-structure model bond prices and actual market bond prices.

As discussed earlier, risk aversion drives a wedge between forward rates and expected future spot rates. Yet risk aversion is not the only factor that separates forward from expected future spot rates. Charts 5 and 6 show separate forward rates and expected future spot rate curves even though the market price of instantaneous interest-rate risk is zero. This type of discrepancy was first explicated by Stanley Fischer (1975) in considering the effect of inflation on nominal interest rates and then by Scott F. Richard (1978) in his two-factor term-structure model. Terence C. Langetieg (1980) observed that only in a world of certainty, not just risk neutrality, would risk premia be zero.

Cox, Ingersoll, and Ross (1981) gave a comprehensive analysis of the phenomenon, which arises from a mathematical condition known as Jensen's inequality.[16] A nontechnical explanation is that equality of forward and future expected spot rates implies a particular type of bond pricing equilibrium, one that is not generally compatible with other types. If, for example, the local expectations hypothesis is true and therefore a bond's current price is the expected discounted value of its face value, where the discount factor is a function of all the instantaneous spot rates expected to prevail up to the maturity date, then forward rates cannot equal expected spot rates in equilibrium. In fact, the bond yield implied by equality of forward and expected spot rates is greater than the yield implied by the local expectations hypothesis.

In other words, even for a zero market price of spread risk, forward rates are biased predictors of future expected spot rates for a reason that has nothing to do with risk premia. However, Charts 5 and 6 make it apparent that the discrepancy between forward and expected future spot rates is very small—only a few basis points (based on re-

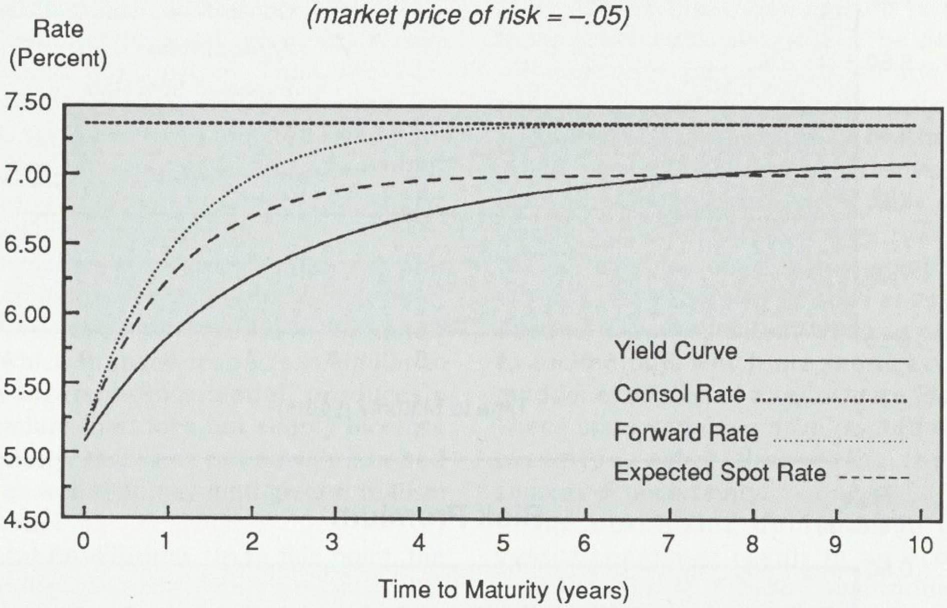
Chart 6.
Cox-Ingersoll-Ross Model
(market pice of risk = 0)



alistic parameter values for the Cox-Ingersoll-Ross model).[17]

Charts 7 and 8 are based on the same model generating Charts 5 and 6, but the market price of instantaneous interest-rate risk is now negative, giving rise to positive risk premia. The expected spot-rate curve is the same as before; however, the forward-rate curve now rises above its level in the earlier charts because of the effect of risk

Chart 7.
Cox-Ingersoll-Ross Model
(market price of risk = –.05)



Risk Premium

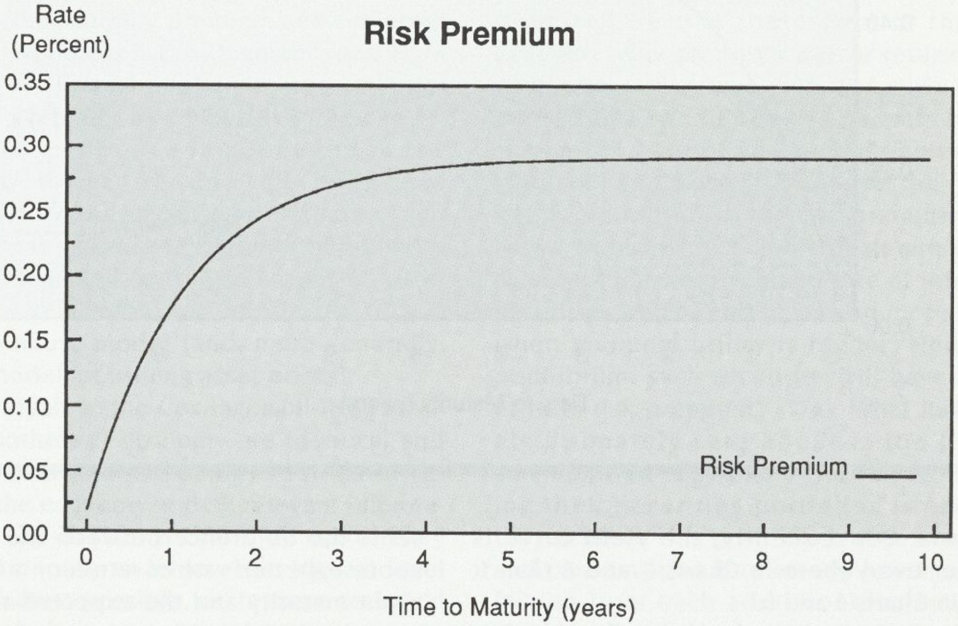
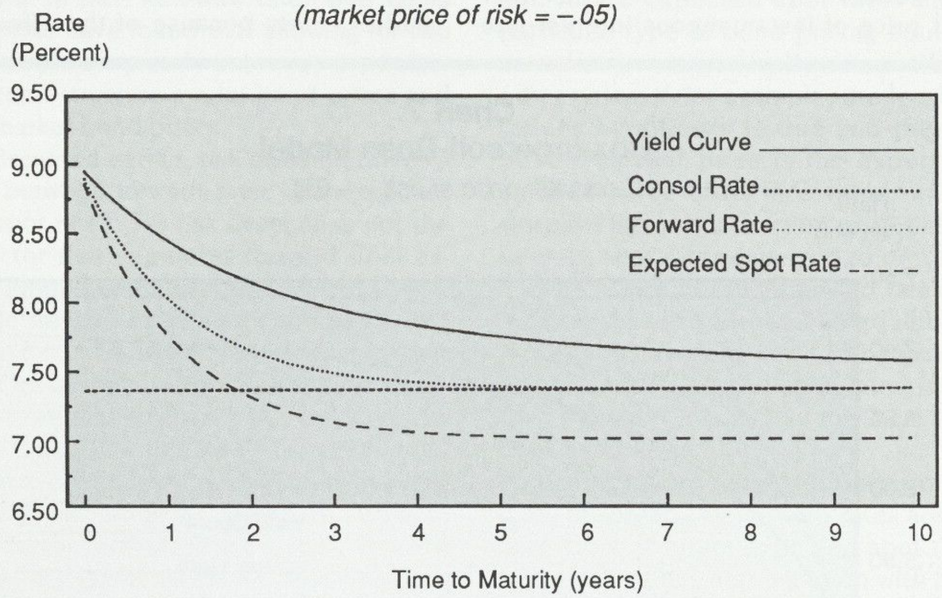
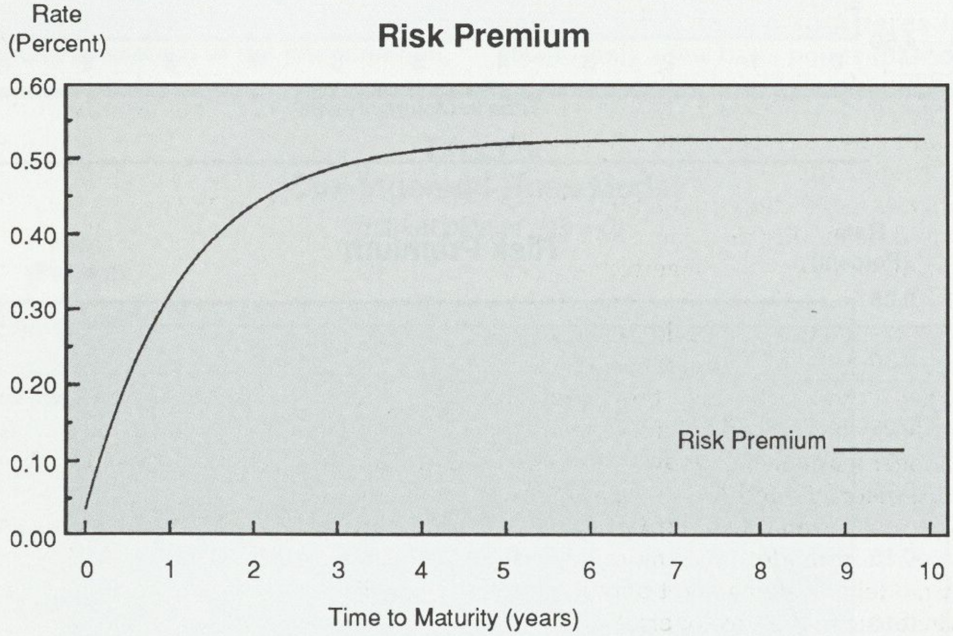


Chart 8.
Cox-Ingersoll-Ross Model
(market price of risk = –.05)



Risk Premium



premia. Consequently, the yield curve is higher everywhere in Charts 7 and 8 than it was in Charts 5 and 6.

The risk premium graphed in the inset below the yield curves in Charts 7 and 8 represents the difference between the instantaneous expected rate of return on a bond of a given maturity and the expected rate of return on a bond that is an instant from maturing (the spot interest rate). This is the risk premium concept used by Cox, Ingersoll, and Ross, which they refer to as the term premium. For both upward- and downward-sloping yield curves, the risk premium curve increases with the maturity of the bond. In fact, for this model the risk premium is proportional to the interest elasticity of the bond price. That elasticity rises with the maturity of the bond. In other words, a 1 percent increase (or decrease) in interest rates depresses (or elevates) longer-term bond prices more than shorter-term bond prices. Thus, the Cox-Ingersoll-Ross model gives a risk premium structure similar to that predicted by the liquidity premium hypothesis. The actual interpretation is different, however, and will be discussed in the next section.

Other models give more complex risk premium behavior. For example, a one-factor model proposed by Francis A. Longstaff (1989), which in many respects is similar to the Cox-Ingersoll-Ross model, produces a risk premium that does not simply increase with maturity; that is, it can have a humped pattern as well that may lend greater realism to the model.

General Equilibrium. Up to this point, the risk premium has been simply a quantity that separates forward rates from expected spot rates. Neglecting the effect of Jensen's inequality, the risk premium's impact is the same in the liquidity premium and preferred habitat hypotheses. Cox, Ingersoll, and Ross (1985a, b) allowed a fuller interpretation of the risk premium by embedding their term-structure model in a larger model of the economy. Particularly since the 1970s, economic theory has placed greater emphasis on the rational maximizing behavior of individuals; Cox, Ingersoll, and Ross sought to incorporate such behavior explicitly into a term-structure model (and, more generally, in any model for valuing asset prices).

Individuals in the Cox-Ingersoll-Ross model's hypothetical economy are identical and rational in the sense that they know the structure of the economy and all relevant information for the decisions that they make. They seek continually to achieve the highest level of satisfaction possible, or, in more technical language, they maximize their "utility functions" through time. They are endowed initially with a certain amount of wealth, which they use to invest in the productive processes of the economy, purchase financial assets (which enable them to defer consumption today to a future date), and consume the economy's single output. Whatever output goes unconsumed is invested to produce more future output.

The individual's maximization problem is to decide on the *optimal* amount of each of these activities to undertake. The decisions will depend in part on their attitude toward risk, which is mathematically represented as an element of their utility function. Risk in this hypothetical economy derives partly from shocks to production that change wealth and comsumption over time. A big rise in the price of oil or the occurrence of drought are examples of such shocks. Production is also uncertain because of shocks to technology, which are modeled as the random evolution of a set of state variables. These state variables have no further interpretation, other than being the basic sources of uncertainty.

The optimizing decisions of the investor/consumer result in an economic equilibrium that includes equilibrium asset prices and interest rates. There are many implications of the Cox-Ingersoll-Ross general equilibrium model. The one to be emphasized here is the nature of the risk premium. Any security's rate of return in excess of the risk-free rate depends on its sensitivity to changes in wealth and to changes in each of the state variables. Cox, Ingersoll, and Ross (1985a) give the following interpretation of the risk premium: "Just as we would expect, individuals are willing to accept a lower expected rate of return on securities which tend to pay off more highly when marginal utility is higher. Hence, in equilibrium such securities will have a lower total risk premium" (376). What individuals ultimately care about is the flow of consumption they enjoy over time. Changing their securities portfolios is one way they achieve this goal. When times are lean and consumption is relatively low, the satisfaction from each additional increment to consumption—that is, its marginal utility—is high. (It is assumed that as consumption rises, marginal utility falls, which is to say that additional consumption carries less utility value.) Trading in securities is the method by which individuals smooth consumption (and raise utility). Because they have limited wealth, (in other words, budget constraints), individuals must prefer some securities more than others. In particular, securities that pay most when consumption is low are more valuable than those that pay most when consumption is high. Consequently, securities that yield most in lean times bear higher prices and hence lower risk premia. That is, their return in excess of the risk-free rate will be lower since risk-averse individuals prefer these kinds of securities. In other words, some securities are better at hedging certain kinds of risk than other securities, and the value of such securities will depend on investor preferences.

Risk-neutral consumers/investors do not care about fluctuations in their consumption flows or investment opportunities. As a result, they would not be willing to pay a premium for the hedging characteristics of any security; all securities would earn the risk-free rate in equilibrium. As mentioned above, however, individuals can be risk averse (specifically, by experiencing diminishing marginal utility with respect to consumption or, equivalently, wealth) but not require any risk premia. As Cox, Ingersoll, and Ross (1981) have pointed out, such a combination could occur if changes in the marginal utility of wealth were unaffected by changes in the state variables. Such conditions should be regarded as exceptional, though, and not realistically characteristic of the economy.

The general equilibrium structure provides a more detailed interpretation of a term-structure model. In particular, the one-factor Cox-Ingersoll-Ross model includes the specialized assumption that the risk premium does not depend on wealth but represents the "covariance of changes in the interest rate with percent changes in optimally invested wealth (the 'market portfolio')" (Cox, Ingersoll, and Ross 1985b, 393). In the one-factor model, a single state variable, which is unobservable, summarizes all underlying uncertainty. However, the spot interest rate depends on that state variable; it determines all movements in the spot rate. The state variable represents changes in physical investment opportunities, which in turn account for all variations in wealth. Thus, risk depends on how changes in bond prices determined by the spot rate vary with changes in wealth. Bonds that are better hedges against shifts in investment opportunities (and hence wealth) are more highly valued and carry lower risk premiums. For example, in the one-factor Cox-Ingersoll-Ross model if wealth tends to decline as the spot interest rate rises (the wealth-spot rate covariance is negative), then short-term bonds will carry lower risk premiums than longer-term bonds because they are subject to smaller capital losses. Because long-term bond prices decline more than short-term bond prices in response to a rise in the spot rate, long-term bonds are riskier to include in the wealth portfolio and require a greater risk premium as compensation. Similar reasoning applies to the interpretation of the risk premium arising in more complex multifactor models. This general equilibrium theory of the risk premium has considerably greater economic content than the comparatively vague notion that investors demand risk premia for illiquid bonds.

The general equilibrium approach of the Cox-Ingersoll-Ross model offers further insights into how underlying economic variables, such as shifts in investment opportunities, affect the term structure. Partial equilibrium models cannot reveal these linkages. In fact, Cox, Ingersoll, and Ross (1985b) make the point that partial equilibrium models can be internally inconsistent. First, there may not be any underlying equilibrium consistent with the assumptions concerning the choice of state variables, such as the spot interest rate, and the way they evolve randomly over time. In other words, the two assumptions may be incompatible. Second, arbitrarily selecting a functional form for the risk premium (for example, assuming that it is constant) may produce a model that implies arbitrage opportunities—a fatal modeling flaw.[18]

From a modeling standpoint, the general equilibrium approach is clearly preferable. In terms of actually implementing a bond pricing model, however, the difference between partial and general equilibrium may not be critical. Even if the general equilibrium model is internally consistent, it may misprice actual bonds as much as a partial equilibrium version. Poor predictions can stem from a misspecification of the actual underlying economy if, for instance, the details of the model are simply too far off. On the other hand, a more complex, realistic model might be impossible to use because essential parameters cannot be reliably estimated given limited data about the actual economy. Simpler models may perform better even if they are internally inconsistent.

Extensions and Variations on Term-Structure Models

Cox, Ingersoll, and Ross (1985b) explore some extensions of their basic model. They show how to incorporate more state variables. This addition results in general equilibrium multifactor models that are similar to the earlier multifactor models of Brennan and Schwartz (1979) and of Schaefer and Schwartz (1984). As more factors are added, the mathematical complexity increases considerably, making these models cumbersome. Simple formulas for pricing bonds are not available; rather, computer-intensive numerical methods must generally be used to obtain bond prices.

Longstaff (1989) attempts to get better performance out of a single-factor general equilibrium model by allowing for nonlinear behavior in the underlying state variable and consequently also in the observable spot interest rate. One impact of this nonlinearity is to make the spot rate revert to its long-term level more slowly from above than from below. Greater flexibility and realism are achieved because the possible theoretical shapes of the term structure are more varied. These attributes may model actual variations in the term structure better than the simple Cox-Ingersoll-Ross model.

Cox, Ingersoll, and Ross (1985b) also show how to adapt their approach to value nominal bonds. Their original model does not include the existence of money or inflation. They modify the model by incorporating the aggregate price level as another state variable and obtain a nominal bond-pricing formula. Alternatively, they recast the entire model in terms of nominal variables and derive interest-rate and bond-pricing equations similar in form to the real-variable model. In this revised model, the instantaneous nominal interest rate equals the sum of the instantaneous real interest, the expected instantaneous rate of inflation, and a group of terms that arise from the effects of Jensen's inequality. In other words, the so-called Fisher equation, which states that the nominal rate is the sum of the real rate and expected inflation, does not hold (for the same mathematical reason that forward rates do not equal expected future spot rates even when factor risk is zero). Richard's (1978) is an earlier partial equilibrium model that included expected inflation as one of two state variables.[19]

The Cox-Ingersoll-Ross model has been used in a number of recent applications involving the valuation of interest-rate contingent claims. Krishna Ramaswamy and Suresh M. Sundaresan (1986) value floating-rate instruments, which have interest payments that vary with current market rates, using the Cox-Ingersoll-Ross single-factor model. The Cox-Ingersoll-Ross model was adapted to valuing the cash flows from such instruments. Louis O. Scott (1989) prices default-free interest-rate caps using the Vasicek and Cox-Ingersoll-Ross single-factor models. John Hull and Alan White (forthcoming) apply modified versions of the Vasicek and Cox-Ingersoll-Ross models to price bond-options and interest-rate caps, which are derivative instruments designed to hedge floating-rate instruments. Peter A. Abken (1990) uses the Cox-Ingersoll-Ross single-factor model as a component of a model to price default-risky interest-rate caps. Sundaresan (1989) employs the Cox-Ingersoll-Ross model in his valuation model for interest-rate swaps, which are also hedging tools that have numerous forms and applications. A basic use is to convert fixed-rate interest payments into floating-rate payments (or vice versa). Longstaff (forthcoming) extends the Cox-Ingersoll-Ross single-factor model to price options that have payoffs specified in terms of yields rather than prices. His preliminary empirical work demonstrates that the model predicts actual traded yield option prices with greater accuracy than other existing models. Abken (forthcoming) develops a swap valuation model that allows for default risk on the part of parties participating in a swap. The Longstaff yield option model is a component of his model.

The single-factor Ho and Lee model has been extended into both discrete- and continuous-time multifactor models by David Heath, Robert Jarrow, and Andrew Morton (1987, 1988). Like the Ho and Lee model, the Heath-Jarrow-Morton models use all the information in the current term structure, but they do so in a more sophisticated way, avoiding the deficiencies of the Ho and Lee model. In particular, the Heath-Jarrow-Morton models preclude negative interest rates and, since they are multifactor, do not imply perfectly correlated bond-price movements. The models are specified in terms of the evolution of the forward rate rather than the bond price as in Ho and Lee.

Unlike Ho and Lee's, these models are preference free, a characteristic that the authors cite as a major advantage of their method. Still, "fitting" the models to the actual forward-rate curve is problematic, and inaccuracies at this step become relayed into inaccurate bond-option prices. Heath, Jarrow, and Morton (1989) conducted an empirical test of the earlier Heath-Jarrow-Morton (1987) model and found that despite some problems, the test supports the model's validity.

In a similar vein, Philip H. Dybvig (1988) shows how to fuse the Ho and Lee model (or, more generally, models incorporating the information in the current term structure) with single- or multifactor state variable models. He finds that a relatively simple single-factor hybrid model fits the term structure very well over time.

Robert R. Bliss, Jr., and Ehud I. Ronn (1989a) have also extended and refined the Ho and Lee approach by including a set of state variables that affect movements in forward rates (that is, the perturbation function and the risk-neutral probabilities are functions of the state variables). Their modeling changes effectively prevent the occurrence of negative interest rates as in the Ho and Lee model and allow for a more flexible modeling of interest-rate movements. In Bliss and Ronn (1989b) the authors test their model by valuing actual options on Treasury-bond futures contracts. The results indicate a systematic discrepancy between model predictions of option prices and actual option prices. A possible explanation for the bias (aside from actual options' being mispriced) is that their estimated perturbation functions, based on past observed forward-rate movements, do not sufficiently capture actual movements in forward rates.

Naoki Kishimoto (1989) extends the Ho and Lee model to enable valuation of assets such as stock options, convertible bonds, and junk bonds, for which risk stems not only from future interest-rate changes but also from other factors. The model combines the features of the Ho and Lee term-structure model with the Cox-Ross-Rubinstein method of modeling asset-price risk.

Black, Emanuel Derman, and William Toy (1990) have quite recently developed another variation on the Ho and Lee model. Their model posits a time-varying volatility for the spot interest-rate process and assumes that the spot rate is log-normally distributed, thereby preventing the realization of negative values. Like Ho and Lee's, this model does not give a simple formula for valuing either bonds or interest-rate contingent claims and consequently must be implemented as a computer algorithm.[20]

Contingent Claims Pricing

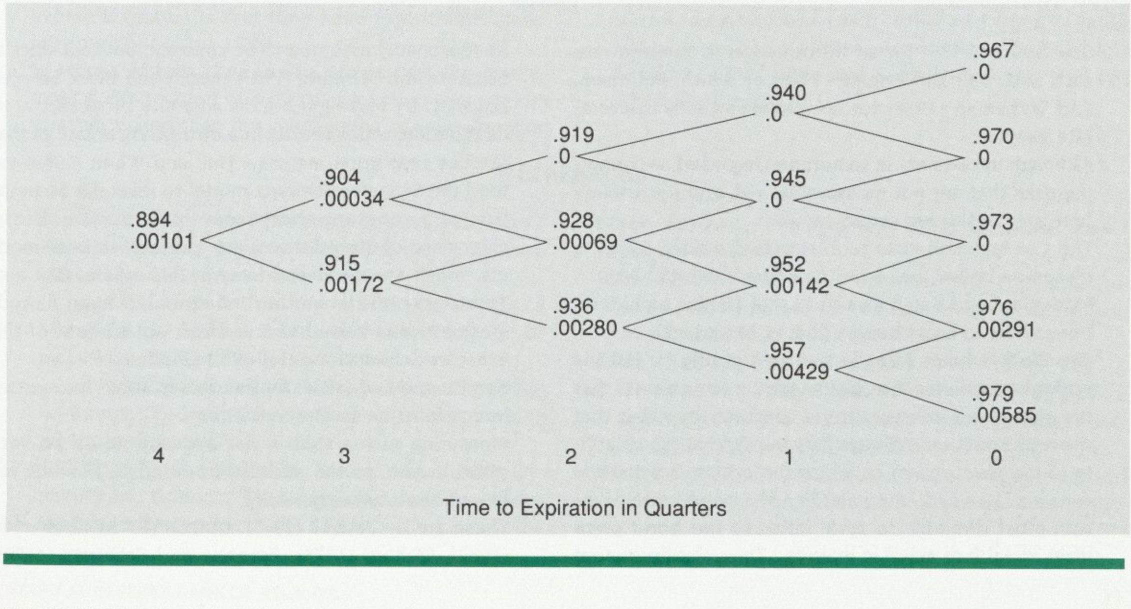
Having introduced the Ho and Lee model, it is now easy to illustrate how an option can be valued using this model and, by extension, any other term-structure model. Consider a European call option on a discount bond—a three-month Treasury bill. A call is usually specified in terms of a bond's price. The option gives the holder the right to buy the T-bill at a prespecified price upon expiration of the option. To make matters easy, the call is assumed to be European so that it may be exercised only upon expiration, unlike an American option, which can also be exercised earlier. The call in the illustration is assumed to have an 11 percent strike or, equivalently, a strike price of 0.973 (on a hypothetical T-bill with $1 face value). The underlying T-bill matures three months after the option. (Note that the option cannot meaningfully expire at the same time as the T-bill because the bill's price is known with certainty, that is, its price equals its face value.)

The object of this exercise is to determine the value of the call option when it has one year remaining before expiration. The earlier example from Chart 3 gives the state-price evolution for the T-bill that is relevant for this problem. Chart 9 repeats the T-bill price tree but also includes the option prices at each node of the tree. Again, the valuation process starts from the expiration date at time 0. The call option has value only when the T-bill price is above the strike price of 0.973 (or equivalently the annualized bill rate is below 11 percent). Thus, the only positive option payoffs occur for the first two lower entries in the column on the far right. The rest of the computation proceeds exactly as it did in Chart 3, except that the call option payoff, not the bill price, is used in recursively working back to the initial date. The probability-weighted call option payoffs in the time 0 column are averaged pairwise and discounted using the three-month bill price corresponding to realized state price one period earlier. That bill price, shown in Chart 3, is derived from Ho and Lee's formula, which in turn is based on the initial term structure of interest rates. Repeating the procedure back to the present time at quarter 4 before expiration gives an option price of 0.00101 on a face value of $1. Multiplying by $1 million to be somewhat more realistic gives a call option price of $1,010. This figure can be converted into an interest rate, expressed in basis points (one-hundredths of a percent), by dividing by 25.[21] Therefore, for every $1 million face value of T-bills, the call option would cost 40.4 basis points.

More accurate valuation would be achieved by dividing the time interval to expiration more finely than into four quarters and thereby increasing the number of branches in the binomial tree. Still other models could be used to value the call option; however, the Ho

Chart 9.
Ho and Lee Model Contingent Claim Valuation

Call Option on Three-Month T-Bill Price (Strike = .973 or 11%)
Value One Year before Option Expiration = $1,010 or 40.4 Basis Points



and Lee model makes the mechanics of valuation fairly transparent.

Cox, Ingersoll, and Ross (1985b) give a closed-form solution for option prices based on their single-factor model.[22] It is similar in form to the Black-Scholes option-pricing formula although more complicated. Farshid Jamshidian (1989) has derived a closed-form solution for pricing bond options based on the single-factor Vasicek model. This formula is much easier to use than the formula of Cox, Ingersoll, and Ross. The choice of which to use is an empirical question that is the subject of current research.

Conclusion

This article has given an introduction to the new types of term-structure models that have appeared in the last fifteen years. These models have evolved not only because of the option-pricing revolution triggered by the Black-Scholes formula and its stimulation of financial research but also as a result of the rational expectations approach in economics and finance. Research in this field has also been motivated by the greater asset-price volatility that has characterized financial markets in recent years. In other words, market demand has been growing for term-structure models to help in the valuation of increasingly complex financial instruments such as interest-rate caps and swaps. It can be expected that financial theory and practice will continue to spur advances in one another.

Much research time and money are currently being devoted to building better term-structure and interest-rate contingent claims pricing models. The benefits of improved modeling have both academic and intellectual significance to financial researchers as well as dollars and cents payoffs to market practitioners. Part of this research effort is devoted to refining existing models to give the best "fit" with actual market bond and option prices. A number of researchers have explored the performance and accuracy of some of the new term-structure models.[23] It is beyond the scope of this article to review their evaluations, but it is important to note that there tend to be trade-offs between ease of use and model accuracy. The type of application may also determine which model is appropriate: a simple single-factor model, for example, may be quite adequate for pricing short-term bond options, while more elaborate multifactor models may be needed for accurate pricing of longer-term options, which require greater precision in term-structure modeling.

Notes

[1] See Abken (1989) for an introduction to interest-rate caps and Wall and Pringle (1988) or Smith, Smithson, and Wakeman (1988) for an introduction to interest-rate swaps.

[2] Although uncertainty is sometimes regarded as connoting risks that are not measurable, risk and uncertainty here are considered synonymous.

[3] The Cox-Ingersoll-Ross term-structure model, which is discussed below, was used to generate these charts.

[4] Modigliani and Sutch also note that similar considerations apply to bond issuers (that is, borrowers).

[5] The Black-Scholes formula performs poorly for valuing options on bonds. The original model assumes (1) that the short-term interest rate is constant (implying that the term structure is always flat) and (2) that the volatility of the asset return on which the option is written is constant. However, the volatility of a bond's rate of return must diminish to zero, because the bond price must equal face value at maturity. The volatility cannot be reasonably assumed to be constant, unlike a stock's. Schaefer and Schwartz (1987) modify the Black-Scholes equation by making a bond's volatility proportional tc its duration, which results in a decreasing volatility that reaches zero upon maturity. Hull and White (1989) extend the Schaefer-Schwartz model to relax the assumption of a constant short-term interest rate. These extensions of Black-Scholes are "preference-free" models, which are discussed later in this article. Ball and Torous's (1983) is another influential, though flawed, model in this vein that is a direct antecedent of the Schaefer-Schwartz and Hull-White models.

[6] See Ho and Lee (1986) for the details about the perturbation function and its derivation.

[7] Measuring such a shift in risk aversion would be very difficult, and, to the author's knowledge, has not yet been attempted empirically

[8] These are by Vasicek (1977), Dothan (1978), Courtadon (1982), Richard (1978), Brennan and Schwartz (1979),

and Langetieg (1980). The Cox-Ingersoll-Ross model first circulated in an unpublished manuscript dated July 1977.

[9]See Black (1988) for Black's surprise and amusement at the continuing popularity of his original model with Scholes. He believes that the original model's highly restrictive assumptions limit its usefulness and suggests modifications to make the model perform better.

The Black-Scholes formula can also be derived using discrete-time mathematics. See Cox, Ross, and Rubinstein (1979).

[10]See Dybvig (1988) for an analysis and extension of the Ho and Lee model.

[11]More precisely, "continuous price dynamics" means that the underlying asset price is assumed to follow an Ito process. See Jarrow and Rudd (1983) for an introduction to option pricing.

[12]See Jarrow and Rudd (1983, chap. 7) for a more detailed exposition of the risk-neutrality argument.

[13]The standard deviations in the denominators are actually negative quantities. The bond return standard deviation can be shown to be the product of the negative interest elasticity of the bond price times the positive standard deviation of the spot interest rate. Therefore positive excess return to volatility ratios (and positive risk premia in the term strucure) occur for negative values of $\lambda(r,t)$.

[14]Ingersoll (1987, chaps. 17-18) gives a succinct, though technical, discussion of valuation methods when a model's state variables are not traded assets. See Cox, Ingersoll, and Ross (1981, 772) for the general method of forming a hedge portfolio for their multifactor term-structure model.

[15]For risk-neutral investors, risk premia are zero only in special cases. Cox, Ingersoll, and Ross (1981) show that one case is the trivial one where interest rates are nonrandom (782-83). Another special case requires that the covariance of the bond's return with the marginal utility of wealth is zero. In general this covariance will not be zero and risk premia will exist. These authors also show that a sufficient condition for the local-expectations hypothesis to prevail for risk-averse investors is that they have logarithmic, state-independent utility (783).

[16]In words, Jensen's inequality states that the mathematical expectation of a concave function of a random variable is less than the value of the concave function evaluated at the mean of the random variable. The inequality relationship is reversed for a convex function. See Cox, Ingersoll, and Ross (1981, 776-77) for an analysis of the effect of Jensen's inequality on different forms of the expectations hypothesis. See Fischer (1975, 513) for a simple example of Jensen's inequality that is also pertinent to the term-structure case.

[17]Campbell (1986) argues that the discrepancy between forward and expected future spot rates is mathematically a second-order effect and can be ignored in doing empirical work. Using his linearized framework for analysis of the term structure, the differences among term structure theories studied by Cox, Ingersoll, and Ross (1981) disappear.

[18]See Cox, Ingersoll, and Ross (1985b, 398) for a simple example of how an arbitrage opportunity could arise.

[19]As a future research topic, it might be interesting to extend Richard's model to allow for the impact of monetary policy on the term structure. For example, anticipated inflation could be specified as a function of current or lagged short-term interest rates. This or some other type of reaction function could probably be incorporated as a first step in modeling the effects of monetary policy.

[20]See Hull and White (forthcoming) for the continuous-time formulation of both the Ho and Lee model and the Black-Derman-Toy model. Dybvig (1988) also gives a continuous-time version of Ho and Lee.

[21]The value of a basis point is derived first by dividing \$1 million by 10,000 (the number of basis points in 100 percent) to give \$100 per basis point per year. The underlying asset, the T-bill, has a three-month maturity; therefore, the value of a basis point per quarter is \$25.

[22]A closed-form solution is a formula that is "ready" to evaluate, given the model's parameters. In contrast, many models, such as Ho and Lee or Brennan and Schwartz, require numerical methods to arrive at bond or contingent claim prices.

[23]This work includes that of Marsh and Rosenfeld (1983), Brown and Dybvig (1986), Dietrich-Campbell and Schwartz (1986), Dybvig (1988), Longstaff (1989), Heath, Jarrow, and Morton (1989), Bliss and Ronn (1989a), Pearson and Sun (1989), Buser, Hendershott, and Sanders (1990), Chen and Scott (1990), and Hull and White (forthcoming).

References

Abken, Peter A. "Interest-Rate Caps, Collars, and Floors." Federal Reserve Bank of Atlanta *Economic Review* 6 (November/December 1989): 2-24.

______. "Valuing Default-Risky Interest-Rate Caps: A Monte Carlo Approach." Federal Reserve Bank of Atlanta Working Paper 90-5, July 1990.

______. "Valuation of Default-Risky Swaps and Options on Swaps." Federal Reserve Bank of Atlanta working paper, forthcoming.

Ball, Clifford A., and Walter N. Torous. "Bond Price Dynamics and Options." *Journal of Financial and Quantitative Analysis* 18 (December 1983): 517-31.

Black, Fischer. "The Holes in Black-Scholes." *Risk* 1 (March 1988): 30-33.

Black, Fischer, and Myron Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81 (May/June 1973): 637-59.

Black, Fischer, Emanuel Derman, and William Toy. "A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options." *Financial Analysts Journal* (January/February 1990): 33-39.

Bliss, Robert R., Jr., and Ehud I. Ronn. "Arbitrage-Based Estimation of Nonstationary Shifts in the Term Structure of Interest Rates." *Journal of Finance* 44 (July 1989a): 591-610.

______. "A Non-Stationary Trinomial Model for the Valuation of Options on Treasury Bond Futures Contracts." University of Texas at Austin unpublished manuscript, November 1989b.

Bodie, Zvi, Alex Kane, and Alan J. Marcus. *Investments*. Homewood, Ill.: Richard D. Irwin, Inc., 1989.

Brennan, Michael J., and Eduardo S. Schwartz. "A Continuous Time Approach to the Pricing of Bonds." *Journal of Banking and Finance* 3 (July 1979): 133-55.

Brown, Stephen J., and Philip H. Dybvig. "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates." *Journal of Finance* 41 (July 1986): 617-30.

Buser, Stephen A., Patric H. Hendershott, and Anthony B. Sanders. "Determinants of the Value of Call Options on Default-Free Bonds." *Journal of Business* 63 (January 1990): S33-S50.

Campbell, John Y. "A Defense of Traditional Hypotheses about the Term Structure of Interest Rates." *Journal of Finance* 41 (March 1986): 183-93.

Chen, Ren-Raw, and Louis Scott. "Maximum Likelihood Estimation for a Multi-Factor Equilibrium Model of the Term Structure of Interest Rates." University of Georgia unpublished manuscript, April 1990.

Courtadon, Georges. "The Pricing of Options on Default-Free Bonds." *Journal of Financial and Quantitative Analysis* 17 (March 1982): 75-100.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross. "A Re-Examination of Traditional Hypotheses about the Term Structure of Interest Rates." *Journal of Finance* 36 (September 1981): 769-99.

______. "An Intertemporal General Equilibrium Model of Asset Prices." *Econometrica* 53 (March 1985a): 363-84.

______. "A Theory of the Term Structure of Interest Rates." *Econometrica* 53 (March 1985b): 385-408.

Cox, John C., and Stephen A. Ross. "The Valuation of Options for Alternative Stochastic Processes." *Journal of Financial Economics* 3 (January 1976): 145-66.

Cox, John C., Stephen A. Ross, and Mark Rubinstein. "Option Pricing: A Simplified Approach." *Journal of Financial Economics* 7 (September 1979): 229-63.

Dietrich-Campbell, Bruce, and Eduardo Schwartz. "Valuing Debt Options." *Journal of Financial Economics* 16 (July 1986): 321-43.

Dothan, L. Uri. "On the Term Structure of Interest Rates." *Journal of Financial Economics* 6 (January 1978): 59-69.

Dothan, Michael U. *Prices in Financial Markets*. New York: Oxford University Press, 1990.

Dybvig, Philip H. "Bond and Bond Option Pricing Based on the Current Term Structure." Washington University in St. Louis Working Paper, September 1988.

Fischer, Stanley. "The Demand for Index Bonds." *Journal of Political Economy* 83 (June 1975): 509-34.

Heath, David, Robert Jarrow, and Andrew Morton. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation." Cornell University Working Paper, October 1987.

______. "Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation." Cornell University Working Paper, April 1988.

______. "Contingent Claim Valuation with a Random Evolution of Interest Rates." Cornell University Working Paper, November 1989.

Ho, Thomas S.Y., and Sang-Bin Lee. "Term Structure Movements and Pricing Interest Rate Contingent Claims." *Journal of Finance* 41 (December 1986): 1011-29.

Hull, John, and Alan White. "An Extension to the Schaefer-Schwartz Bond Option Pricing Model." University of Toronto unpublished manuscript, February 1989.

______. "Pricing Interest-Rate Derivative Securities." *Review of Financial Studies* 3 (forthcoming).

Ingersoll, Jonathan E., Jr. *Theory of Financial Decision Making*. Totowa, N.J.: Rowman and Littlefield, 1987.

Jamshidian, Farshid. "An Exact Bond Option Formula." *Journal of Finance* 44 (March 1989): 205-9.

Jarrow, Robert A., and Andrew Rudd. *Option Pricing*. Homewood, Ill.: Richard D. Irwin, Inc., 1983.

Kishimoto, Naoki. "Pricing Contingent Claims under Interest Rate and Asset Price Risk." *Journal of Finance* 44 (July 1989): 571-89.

Langetieg, Terence C. "A Multivariate Model of the Term Structure." *Journal of Finance* 35 (March 1980): 71-97.

Longstaff, Francis A. "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates." *Journal of Financial Economics* 23 (August 1989): 195-224.

______. "The Valuation of Options on Yields." *Journal of Financial Economics* (forthcoming).

Marsh, Terry A., and Eric R. Rosenfeld. "Stochastic Processes for Interest Rates and Equilibrium Bond Prices." *Journal of Finance* 38 (May 1983): 635-45.

Modigliani, Franco, and Richard Sutch. "Innovations in Interest Rate Policy." *The American Economic Review* 56 (May 1966): 178-97.

Pearson, Neil D., and Tong-Sheng Sun. "A Test of the Cox, Ingersoll, Ross Model of the Term Structure of Interest Rates Using the Method of Maximum Likelihood." Unpublished manuscript, January 1989.

Ramaswamy, Krishna, and Suresh M. Sundaresan. "The Valuation of Floating-Rate Instruments: Theory and Evidence." *Journal of Financial Economics* 17 (December 1986): 251-72.

Richard, Scott F. "An Arbitrage Model of the Term Structure of Interest Rates." *Journal of Financial Economics* 6 (March 1978): 33-57.

Ritchken, Peter, and Kiekie Boenawan. "On Arbitrage-Free Pricing of Interest Rate Contingent Claims." *Journal of Finance* 45 (March 1990): 259-64.

Schaefer, Stephen M., and Eduardo S. Schwartz. "A Two-Factor Model of the Term Structure: An Approximate Solution." *Journal of Financial and Quantitative Analysis* 19 (December 1984): 413-24.

______. "Time-Dependent Variance and the Pricing of Bond Options." *Journal of Finance* 42 (December 1987): 1113-28.

Scott, Louis O. "Pricing Floating Rate Debt and Related Interest Rate Options." University of Georgia unpublished paper, October 1989.

Smith, Clifford W., Charles W. Smithson, and Lee Macdonald Wakeman. "The Market for Interest Rate Swaps." *Financial Management* 17 (Winter 1988): 34-44.

Sundaresan, Suresh. "Valuation of Swaps." Columbia University Center for the Study of Futures Markets Working Paper 183, April 1989.

Vasicek, Oldrich A. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics* 5 (November 1977): 177-88.

Wall, Larry D., and John J. Pringle. "Interest Rate Swaps: A Review of the Issues." Federal Reserve Bank of Atlanta *Economic Review* 73 (November/December 1988): 22-37.

Recent Views of Viewing the Real Rate of Interest

Rosemary Thomas Cunningham and
Thomas J. Cunningham

$$rr_{ep} - rr_{ea} = (i - \pi^a) - (i - \pi^e) = \pi^e - \pi^a = \text{forecast error.}$$

The question of whether monetary authorities have any influence over real interest rates has had economists' attention for a long time. The authors review various theoretical approaches to this issue and recent evidence in this article. They include a report on their own recent research based on a small, opening economy. Consistent with some earlier empirical work, their findings suggest an avenue of influence for the monetary authority over real interest rates.

Interest rates play a pivotal role in determining economic activity, and thus the behavior of the real, or inflation-adjusted, rate of interest is central to understanding macroeconomic dynamics. Investment expenditures, consumer durable expenditures, and saving are particularly sensitive to movements in real rates of interest.

This article is an attempt to understand recent developments in economic theory relating to the real rate of interest and, in particular, the ability of the monetary authority to influence it. Several perspectives deserve attention. If real processes and preferences within the economy give rise to the real rate, the real rate then conveys information about the economy's preference for consumption now rather than in the future (that is, it represents the intertemporal marginal rate of substitution). If the real rate is thought of as the marginal product of capital, it provides information about the potential profitability of investment. In either case the monetary authority has no influence over the real rate process. If the real rate is viewed as being the outcome of a market for loanable funds, however, monetary policy does gain some influence over its behavior. Finally, if the economy is small relative to the rest of the world and open to international capital flows, the real rate may be influenced by, and indicate something about, relevant world conditions.

The authors are, respectively, an assistant professor of economics at Agnes Scott College and a senior economist in the macropolicy section of the Atlanta Fed's research department.

Understanding the real rate of interest has been quite problematic ever since Irving Fisher (1930) made the distinction between the ex ante and the ex post real rate of interest. Because the ex post real rate of interest is simply the nominal rate of interest less the rate of inflation, it can be observed. In contrast, the ex ante real rate cannot be observed because it is the nominal rate of interest minus the expected rate of inflation. The ex ante real rate is the one that influences economic decisions and should be thought of as the real rate of interest even with the obvious problem it presents.

A recent solution to the dilemma posed by the ex ante real rate's being unobservable focuses on the problem of expectation formation and imposes some restrictions on that process in order to approximate an ex ante real rate of interest series. This article provides an overview of the theory behind the problem, followed by a discussion of the proposed solution and the implications of the restrictions it imposes. The concluding sections present applications of the solution, including recent research based on a small, opening economy, and summarize the role of the monetary authority in relation to real interest rates.

Economic Theory and the Real Rate of Interest

In Fisher's (1930) division of nominal interest rates, he claimed that the nominal (observed) rate of interest, i, was composed of the real rate of interest, rr; an adjustment for the inflation (denoted by π) that was expected to occur over the relevant time period, π^e; and the cross product of the two, $\pi^e \times rr$. Thus the basic Fisher equation is

$$i = rr + \pi^e + (\pi^e)(rr).$$

Most empiricists ignore the cross-product term since it is quite small under conditions of modest inflation and real rates of interest.[1] The resulting equation for the real rate of interest is

$$rr = i - \pi^e.$$

There are two prevalent misuses of the Fisher equation. First, assuming that the real rate is constant (usually 2 or 3 percent), any change in nominal interest rates could be interpreted as a change in expected inflation. Although many claim that the real rate of interest may be constant over the long run, drawing inferences about expected inflation in short-run analysis frequently results in error.

A more problematic misapplication of Fisher's work is the substitution of *observed* rates of inflation (as in ex post real rates of interest) for *expected* rates (as in ex ante rates) in the construction of a real rate of interest series. Although the two are not interchangeable, the substitution is frequently made because of the problem associated with observing the expected rate of inflation.[2] The difference between the ex post and ex ante rate is the forecast error for inflation during the period in question, and this error can be substantial, as may have been the case of fixed-rate mortgages in the 1960s. During that time, the ex post rate of return on mortgages seemed to be negative, by a large margin, and it seems unlikely that lenders would have made fixed-rate loans on the basis of negative ex ante rates.

Before developing the issue of the substitutability of the ex post for the ex ante real rate and the assumptions and restrictions that the substitution implies, an overview of various theories of the real interest rate's behavior is in order. The theories provide alternative approaches to interpreting movements in the real interest rate and to assessing the effectiveness of policy, particularly monetary policy, in influencing the real rate.

The most familiar class of model wherein monetary policy—the actions of the central bank—has any influence over the real interest rate is the Keynesian "liquidity effect," which views the rate of interest as representing the price of holding on to money. From this perspective, it follows that a relative increase in the money supply should lower interest rates.

A problem with this approach, however, is that the Keynesian paradigm has trouble accounting for inflation. The Keynesian world is one in which prices are essentially fixed. In a fixed-price world—without inflation and thus without any inflation premium—a decrease in "the rate of interest" is equivalent to a decrease in the real interest rate. The idea of fixed prices may be a useful assumption in certain situations but clearly is inappropriate for contemporary thinking about the effects of money supply increases. Because of this shortcoming, the Keynesian model has fallen out of favor for policy analysis.[3]

The monetary authority also has some influence over the real rate of interest in the loanable funds model, according to which the real rate of interest is the outcome of market interaction between savers, who supply funds, and borrowers, who demand funds. The real interest rate is the price that adjusts to clear the market, with aggregate inflation expectations simply added to the real rate outcome.

In this model the monetary authority exerts its influence over the real rate in two ways. First, by conducting open market operations the central bank directly enters the market to purchase or sell securities. Second, to the extent that the central bank can influence the operations of other financial intermediaries, usually through reserve requirements, it can induce others to provide funds to the market, thereby influencing the real rate.

A more prevalent way of thinking about the real interest rate centers on the real side of the economy and equates the real interest rate with the marginal product of capital (the output associated with adding one additional unit of capital to the economy). The argument is that if this marginal product of capital is greater than the real rate of interest, borrowing the money to add to the capital stock is profitable. The agent making the investment must pay the real rate of interest on the debt created to purchase the capital but receives the additional output, the real return, from that added capital. The difference is profit to the investing agent. As long as the profits are positive—that is, as long as the real interest rate is less than the real return on that investment (the marginal product of capital)—additional investment will continue because it is profitable.

As the additional investment occurs, two things happen. First, the demand for capital bids up the real interest rate. At the same time, the additional investment drives down the marginal product of capital, decreasing the profitability of added capital. The investment process ends when real profitability reaches zero (that is, when the marginal product of capital equals the real rate of interest).

The marginal product of capital model has gained popularity because of its explanatory power. It results from general equilibrium models, which attempt to describe an entire economy and can portray the interaction among various markets and potential transmission of disturbances from one sector to the rest of an economic system. For example, any disturbance to individuals' savings preferences will show up as disturbances in the real interest rate and thus in the amount of investment undertaken in the aggregate economy. Similarly, technological innovations that raise the marginal product of capital would stimulate investment, and disturbances to the supply of production inputs (oil supply shocks, for example) would alter the productivity of capital, affecting investment and the real rate.

"A more prevalent way of thinking about the real interest rate . . . equates the real interest rate with the marginal product of capital (the output associated with adding one additional unit of capital to the economy)."

In the marginal product of capital model only real factors influence the real rate of interest. Purely nominal disturbances have no effect. In Keynesian terminology, it is as if the economy suffered no "money illusion." What matters is not money prices but real prices, as though everyone in the economy could immediately understand the distinction between general changes in the price level—simple inflation—and relative changes in the prices of goods, representing some real difference in the economy.

Modifications to the marginal product of capital model give the monetary authority some influence, albeit temporary, over the real rate of interest by assuming that economic agents cannot make the above distinction correctly and instantaneously. During the period in which economic agents are sorting out relative price changes from general price level changes, the monetary authority may influence the real rate. Ultimately, when agents understand the changes and have full and complete information, the monetary authority loses control over any real facet of the economy and has influence over only the nominal price level.

"Modifications to the marginal product of capital model give the monetary authority some influence, albeit temporary, over the real rate of interest by assuming that economic agents cannot make the above distinction correctly and instantaneously."

The central theme of this "policy ineffectiveness" result is that the monetary authority has no influence over real activity except what these informational disparities make possible.[4] Additional assumptions about how the economy learns over time could make these informational disparities quite short-lived and reduce the monetary authority's control to that of the inflation rate only.

If in fact the monetary authority influences only the inflation rate, the Fisher equation indicates that there will be a somewhat perverse effect on the observed nominal interest rates: a policy of monetary ease (supplying relatively more money) would result in higher nominal interest rates as the markets translate the faster growth rate of money into higher expected inflation.

In principle, the difference between this model and the Keynesian model can be tested empirically. Thomas J. Cunningham (1987) described the difficulties of detecting liquidity effects and concluded that models allowing the monetary authority its perverse influence on interest rates are generally the most appropriate for describing the economy; that is, it is largely accepted that the real rate of interest is beyond the influence, even in the short run, of the monetary authority. Some recent evidence challenges this view, however, and will be discussed below.

An alternative way to view the real interest rate in a general equilibrium setting is to envision the real return on saving as a reward for abstaining from consumption. Imagine that at the end of the production cycle there exists a set of goods that can either be consumed directly or invested in the next production cycle. Firms want the output to be invested productively and are willing to offer a real reward to consumers willing to forgo some consumption. The reward is the real rate of interest. As discussed earlier, firms will be willing to offer a real return up to the value of the marginal product of capital.

The consumer, on the other hand, must weigh the value of consumption now against its value later. Because future consumption is less desirable to consumers, they must be induced to delay some consumption by being given the potential for greater consumption in the future—the real return on savings.

From the consumption standpoint, this real rate of interest represents consumers' rate of time preference for consumption now versus consumption later. If the real interest rate is higher than a consumer's rate of time preference, the consumer will save more because the ultimate reward for saving is considered relatively large. By the same token, if the real interest rate is less than the rate of time preference, the consumer will save less.

Ultimately the real rate of interest must coincide with the economy's aggregate rate of time preference—a real factor over which the monetary authority has no effect. Taken with the demand for capital discussed above, the real interest rate grows out of the interaction between the marginal product of capital and the aggregate rate of time preference, neither of which the central bank influences. The real rate of interest is, according to the logic of this model, exogenous to monetary policy.

Thus far the discussion of the behavior of real interest rates has assumed an economy that is not open to international capital flows. Particularly as the economies of the world are becoming increasingly integrated, this assumption can be misleading. The implications of an open economy for monetary policy and the real rate of interest hinge on the type of exchange rate regime to which the monetary authority commits itself: a fixed exchange rate, tying the value of the currency to that of another currency, a basket of currencies, or some commodity; or a flexible exchange rate, whereby market forces determine the international value of domestic money.

Operating under a fixed exchange-rate system, the monetary authority will use monetary policy to stabilize the currency's international value.[5] Committed to this exchange-rate regime, monetary policy cannot deviate from a course of keeping exchange rates constant. Consequently, many questions concerning the influence of the monetary authority become moot. Problems of the domestic economy's price level stability and movements in the real interest rate are beyond the immediate concern of the monetary authority and are resolved by interaction with the rest of the world. Any difference in the domestic real rate of interest from the world real rate would immediately be arbitraged away by a sufficiently large flow of capital toward the relatively higher return.

Flexible exchange rates, on the other hand, allow monetary authorities to pursue a monetary policy independent of pure exchange-rate considerations. Because the monetary authority does not have to conduct policy in a manner exclusively designed to stabilize the exchange rate, it is free to pursue other goals. Indeed, a major impetus for the breakdown of the post–World War II Bretton Woods agreement (establishing a fixed-rate regime) was the difficulty of coordination among the major central banks with differing policy goals.

Under a flexible exchange-rate system, currencies are traded on an open market, much as any other financial instrument. Relative supply and demand will determine a currency's international value. Interest rates can differ among countries without engendering a large capital flow because the international return on domestic investments is no longer simply the domestic interest rate but the domestic interest rate plus the change in the currency's international value during the period the investment is held. Thus any difference in real returns between domestic economies can be equalized by a real change in the relative value of the currencies.

In the framework of the models discussed above, opening the economy to international financial markets simply shifts the level of discussion from that of the nation to that of the world. World real rates of return tend to equalize, and, if a nation's economy is sufficiently small and open relative to the rest of the world, its adjustment process can be quite quick. At the extreme, the small open economy has no ability to control independently its domestic real rate of interest, price level, and exchange rate. The three must simultaneously adjust to equate (among other things) real rates of return with the rest of the world.

With the exception of the work discussed below, there is little evidence that the monetary authority can influence real interest rates. Indeed, what evidence there is to support the notion that the monetary authority has any influence over even nominal interest rates suggests that the effect is to raise, not lower, nominal interest rates; an increase in the money supply (or money supply growth rates) will lead to an increase in nominal interest rates so that an "easy money" policy actually is associated with higher, not lower, nominal interest rates.

Two hypotheses have been put forth to explain this result. The "policy anticipations" or "expected liquidity" hypothesis suggests that a liquidity effect does exist but that an unexpected increase in the money supply today leads to expectations of lower money growth rates in the future; nominal interest rates rise because of an expected liquidity effect in the future. This is a popular explanation for results during the period from 1979 to 1982, when the Federal Reserve System had an operating procedure focused on targeting monetary growth rates, thus giving the central bank a motive to ensure that deviations in the expected path of money growth were eventually offset. A second hypothesis, the inflation expectations hypothesis, argues that an increase in current money supply growth causes expectations that inflation will rise, thereby pushing up nominal interest rates according to a straightforward Fisher effect.[6]

Two important conclusions follow from the foregoing empirical analysis. First, there is substantial evidence that an easing of monetary policy will lead to higher nominal interest rates. The second, and more important, point is that monetary policy cannot claim to influence the real rate of interest. The problem, however, discussed above, is that the relevant measure of real rates of interest is unobservable and thus out of the reach of simple econometric models. It is to the solution of this problem that attention is now turned.

How to Solve the Problem: An Application of Rational Expectations

Fredric Mishkin of Columbia University has recently exploited rational-expectations restrictions to construct a real interest-rate time series useful for econometric inquiry. "Rational expectations" theory maintains simply that economic agents cannot systematically be fooled in forming expectations; that is, people may be continuously incorrect in their expectations, but the mistakes they make are unpredictable. Stated simply, people learn enough from their mistakes that they do not repeat the same error.

This simple insight turns out to be rather powerful in practice. Consider the real interest rate equations described above. The goal is to be able to derive the ex ante real rate of interest, which, to repeat with added notation, is

$$rr_{ea} = i - \pi^e.$$

However, we are able to observe only the ex post real rate of interest,

$$rr_{ep} = i - \pi^a,$$

where π^a represents the actual rate of inflation. The difference between these two measures of the real rate is simply the inflation forecast error:

$$\begin{aligned} rr_{ep} - rr_{ea} &= (i - \pi^a) - (i - \pi^e) \\ &= \pi^e - \pi^a \\ &= \text{forecast error.} \end{aligned}$$

According to the rational expectations hypothesis, the mathematical expectation of the forecast error is zero. Thus, as long as the notion and implications of rational expectations theory are acceptable, the observable ex post real rate of interest series can be directly substituted for the unobservable, but economically important, real rate of interest.

Applications of the Solution

Mishkin has used rational expectations to investigate, among other things, the effects of monetary policy changes on real interest rates, the economic effects of high real interest rates in the United States during the 1980s, and the equality of real interest rates across countries. The approach and work is relatively new. While the rational expectations restrictions discussed above seem rather simple and innocuous, the notion of empirically estimating real rates of interest is quite controversial.

Mishkin's work with John Huizinga (1986) indicates that monetary policy has been the source of the high real interest rates of the 1980s in the United States. In particular, the researchers conclude that the high real interest rates of that decade were the result of the regime change adopted by the Fed in order to disinflate the U.S. economy.

Huizinga and Mishkin focus not only on the high level of real interest rates in the early 1980s and its relation to the Fed's policy changes of October 1979 and October 1982 but also on what they term the significant shift in the stochastic process of real interest rates—the behavior of the real rate series as it bounces around over time.

In October 1979 the Federal Reserve System changed its policy from one of targeting interest rates to targeting nonborrowed reserves. In October 1982 the Fed switched to a borrowed reserve target. Huizinga and Mishkin explored the possibility that, after the policy changes of 1979 increased uncertainty about interest rates, money supply growth and inflation played a role in real interest-rate movements. In addition, they investigated a similar regime change that took place in 1920 when, in response to high inflation rates, the Fed dramatically increased the discount rate twice.

Using data from 1953 to 1984, Huizinga and Mishkin employ rational expectations theory to describe the ex ante real rate of interest as a linear function of a constant, the one-month nominal interest rate, the one-month inflation rate lagged one and two months, and a supply-shock variable (measured as the logarithm of the relative price of fuel and related products in the producer price index) lagged one month. In substituting the ex post real rate for the ex ante real rate, they substituted the nominal interest rate on the one-month bill less the rate of inflation calculated from the consumer price index.

To test whether or not monetary policy regime changes influence the manner in which these variables affect the real rate of interest, Huizinga and Mishkin tested for the equality of the coefficients during various time periods. They found that a shift did indeed occur in these coefficients in October 1979 and October 1982. While in all sample periods a significant negative relationship exists between real rates and expected inflation, the relationship between real and nominal interest rates in the 1979-82 period differs from those in the pre-1979 and the post-1982 periods. In the 1953-79 and 1982-84 periods, there is a significant negative relationship between real and nominal interest rates, whereas in the 1979-82 period the relationship is significantly positive.

Huizinga and Mishkin hypothesized that if a break occurred during the early 1980s in the stochastic process of real interest rates, there should also be breaks in the process during other periods of monetary policy regime changes. Although the exact data series used for the latter time periods were not available in the 1916-27 period, Huizinga and Mishkin worked using comparable data and again tested for the equality of coefficients before and after the monetary policy change in 1920. The results of their tests indicate that, similar to the effect of the monetary policy regime changes of the early 1980s, there was a significant shift in the stochastic process of real interest rates.

Mishkin (1988) explains that because there is not one real rate of interest, monetary policy may affect various sectors of the economy differently. In particular, he claims that during the 1980s agriculture bore a disproportionate share of the disinflationary burden. This situation occurred not only because farmers depended heavily on debt markets in financing production but also because real interest rates were much higher in that sector than in the economy as a whole.

According to Mishkin the relevant real interest rate for the agricultural sector is not one that should be calculated with a broad-based price index such as the consumer price index but rather with one that reflects the prices of farm products. To do this, he constructed a real rate of interest series by subtracting the rate of increase in the farm products component of the producer price index from the one-month Treasury bill rate for the period from January 1953 to December 1986.

Using the relevant real interest rate for the agricultural sector for the periods from January 1953 to October 1979 and from November 1979 to December 1986, Mishkin then estimated essentially the same equation as in Huizinga and Mishkin (1986). The results indicated that the only significant explanatory variable in either period was the one-month Treasury bill rate.

Mishkin also compares his estimates of real interest rates in the agricultural sector with real interest rates for the economy as a whole. In his plot of the two graphs, the real rate for the farm sector appears to be an amplified movement of the real rate for the overall economy. Mishkin confirms this conclusion by regressing the estimated agricultural real rate with the overall real rate and finds that there is a statistically significant comovement between the two and that the agricultural real rate moves more than one-for-one with the overall real rate. From this information, he concludes that the agriculture sector bore much of the burden of the disinflationary period of the early 1980s.

Using data from the Eurodeposit market for seven industrialized nations, Mishkin (1984) also investigated the equality of real interest rates across countries. Again, if a nation is engaging in free trade and there are no restrictions on the mobility of capital, then real interest rates should be equal across countries. If his theory is correct, domestic monetary authorities have no control over real interest rates and their ability to use expansionary or contractionary monetary policy to influence economic activity is therefore constrained.

Mishkin investigated the equality of real interest rates across the United States, Canada, the United Kingdom, France, West Germany, the Netherlands, and Switzerland from the second quarter of 1967 to the second quarter of 1979. He calculated the ex ante real rate of interest two ways: the three-month Eurodeposit rate less the rate of change in the consumer price index and the three-month Eurodeposit rate less the rate of change in the wholesale price index for a particular country. The results indicated that these nations' real interest rates were not equal. Mishkin asserts that real interest rates may differ across countries because risk premiums may differ and because transactions costs and the lack of perfect substitutability of goods implies the violation of purchasing power parity; it does not necessarily imply that real rate arbitrage opportunities exist.

He also asked whether real rates at least moved similarly over time from economy to economy even if they were not strictly equal among nations. The results imply that real rates of interest do not move in similar ways across nations. Mishkin's results indicate that, although capital markets may be integrated, real rates may differ across countries and the domestic monetary authorities may therefore be able to employ monetary policy to influence economic activity.

Real Interest Rates in a Small, Open Economy

We recently began some work that has been in part motivated by Mishkin's work concerning real rates of interest and adopts his restrictions (Cunningham and Cunningham 1990). It centers, though, on the theoretical conformity of the behavior of real rates of interest in a small, open economy. Although real rates of interest may not be equal or move similarly over time across the developed nations that Mishkin studied, the situation may be quite different for a small, open economy with which a relatively large economy such as the United States is a major trading partner.

For a small, open economy the actions of the domestic monetary authorities may determine real rates of interest, or world (U.S.) real interest rates may be the primary influence on real interest rates. The idea that domestic monetary forces control real interest rates simply extends Huizinga and Mishkin's work to small, open economies. The possibility that world real interest rates have the greater influence is in accord with standard open economy macroeconomic theory and, if true, would indicate that U.S. monetary policy has not only domestic effects but also influence over economic activity in at least some other nations.

To investigate the behavior of real interest rates in a small open economy, we looked at the case of the Republic of China on Taiwan from October 1985 through January 1990. Taiwan is an interesting subject because in July 1987 the country moved from a functionally closed economy to one that is essentially open. Prior to financial liberalization, foreign investment was limited to certain types of direct investment in manufacturing, and domestic investment in foreign firms was restricted. In July 1987 reforms were enacted that opened capital markets, relaxing foreign exchange restrictions and permitting international capital flows.

To examine the relative influences of domestic monetary policy and the U.S. real rate of interest on the real interest rate in Taiwan, we regressed the real rate of interest in Taiwan on the growth rate of the domestic money supply, growth in industrial production, and the real rate of interest in the United States. The results could indicate one of three situations. First, the real interest rate in Taiwan could have been affected by the U.S. real rate of interest even before the economy was opened financially because of the fact that the real side of the economy was relatively open, thereby potentially bringing the domestic marginal product of capital in line with the world (U.S.) marginal product of capital. Second, the real rate of interest could have been influenced by the action of the domestic monetary authorities before July 1987 and influenced by U.S. real interest rates thereafter, with the domestic monetary authority's having lost whatever control it had had prior to the financial opening. And third, the primary influence on real interest rates both before and after financial liberalization could have been the domestic monetary authority only.

Our preliminary results are consistent with standard economic theory. They provide some evidence of conditions similar to a liquidity effect and are consistent with Mishkin's work in that, both before and after the financial opening of the economy, the change in the domestic money supply is a significant explanatory variable of real interest rates in Taiwan. The results, however, also indicate that after financial liberalization, U.S. real rates of interest influenced real interest rates in Taiwan.

Overall, our results have intuitive appeal. There is some evidence that before the financial opening of an economy, the domestic monetary authority can influence real rates of interest. After the financial opening, the monetary authority still has some influence, but world real interest rates also become important. The significance of the study is tempered by a somewhat small data set, however, and preliminary results should best be considered simply suggestive. At the same time, the validity of the results is supported by their consistency with standard open-economy macroeconomic theory.

Conclusion

The question of whether the monetary authority may have some influence over real interest rates is far from settled. A substantial body of empirical work shows a direct relationship between money growth rates and nominal interest rates.

Nevertheless, the work discussed above intimates an avenue of influence for the monetary authority over real interest rates. Mishkin's work suggests a liquidity effect (in the real interest rate) for developed economies, and our work extends that result to a newly industrialized economy. Our work also suggests that the overall influence of the monetary authority in a small, open economy may be tempered substantially by real rate movements in the rest of the world. Nonetheless, economies with financial markets in a limited state of development may find the judicious use of monetary policy effective in influencing real rates of interest.

Notes

[1] This article's analysis also ignores tax complications and the term structure. The information about economic expectations contained in the term structure is a separate issue from that discussed here, which relies on the convenient conventional assumption that there is "an" interest rate.

[2] Some surveys are attempting to address this issue and can be expected, eventually, to provide a potentially useful source of data for certain countries during certain times.

[3] An early attempt to solve this problem came, independently, from Robert Mundell and James Tobin. The Mundell-Tobin effect allowed for the incorporation of inflation into the Keynesian IS-LM model, and the effect of this inflation was to reduce marginally the real rate of interest while raising the equilibrium level of income. See Blanchard and Fischer (1989, 546).

[4] The concept is now pervasive enough to show up in undergraduate economics texts.

[5] One of the most frequently forgotten rules of economic policy is the assignment principle, which says that for each and every policy target there must be a policy instrument—that is, for every economic variable to be influenced (inflation, unemployment, exchange rates, and so forth), the policymaker needs to have a policy instrument (such as money growth or government spending) to manipulate. The dilemma facing the monetary authority is that it is limited to using the

policy instrument to target one economic variable at a time. For example, suppose the monetary authority wishes to target a specific rate of inflation. If it adopts a policy of money growth designed to engender the desired effect, it cannot respond to nominal interest rate movements without abandoning the initial policy. Having chosen the rate of inflation as the target, the policymaker must accept whatever happens to other economic variables, including perhaps undesired consequences.

[6] These hypotheses are described in Cunningham and Cunningham (1991), Cunningham and Hardouvelis (1987), and an excellent review by Sheehan (1985) and touched on in Cunningham (1987).

References

Blanchard, Olivier Jean, and Stanley Fischer. *Lectures on Macroeconomics*. Cambridge, Mass.: MIT Press, 1989.

Cumby, Robert E., and Frederic S. Mishkin. "The International Linkage of Real Interest Rates: The European-U.S. Connection." National Bureau of Economic Research Working Paper 2876, 1987.

Cunningham, Thomas J. "Observations about Observations about Liquidity Effects: The Difficulties of Exploring a Simple Idea." Federal Reserve Bank of Atlanta *Economic Review* 72 (September/October 1987): 14-22.

Cunningham, Thomas J., and Rosemary Thomas Cunningham. "The Behavior of Real Rates of Interest in a Small, Opening Economy." Federal Reserve Bank of Atlanta working paper, forthcoming 1990.

______. "Another Source of Money Supply Announcement Surprises." *Journal of Macroeconomics* (forthcoming 1991).

Cunningham, Thomas J., and Gikas A. Hardouvelis. "Temporal Aggregation, Monetary Policy, and Interest Rates." Federal Reserve Bank of Atlanta Working Paper 87-4, May 1987.

Fisher, Irving. *The Theory of Interest*. New York: Macmillan Publishing Company, 1930.

Huizinga, John, and Frederic S. Mishkin. "Inflation and Real Interest Rates on Assets with Different Risk Characteristics." *The Journal of Finance* 39 (July 1984): 699-712.

______. "Monetary Policy Regime Shifts and the Unusual Behavior of Real Interest Rates." National Bureau of Economic Research Working Paper 1678, 1985.

______. "Monetary Policy Regime Shifts and the Unusual Behavior of Real Interest Rates." *Carnegie-Rochester Conference Series on Public Policy* 10 (Spring 1986): 231-74.

Kuo, Shirley W.Y. *The Taiwan Economy in Transition*. Boulder, Co.: Westview Press, 1983.

Mishkin, Frederic S. "Are Real Interest Rates Equal Across Countries? An Empirical Investigation of International Parity Conditions." *The Journal of Finance* 39 (December 1984): 1345-57.

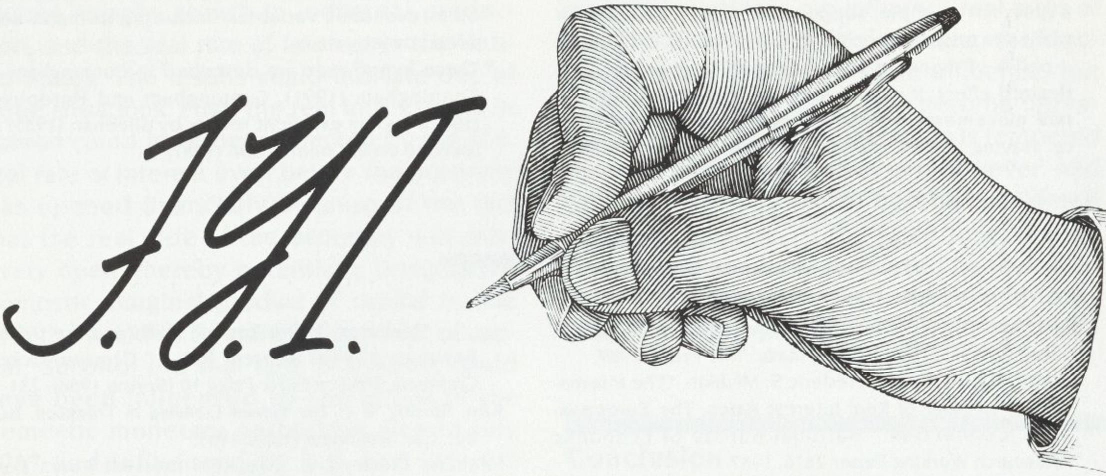
______. "The Real Interest Rate: A Multi-Country Empirical Study." National Bureau of Economic Research Working Paper 591, 1985.

______. "Understanding Real Interest Rates." National Bureau of Economic Research Working Paper 2691, August 1988.

Sheehan, Richard G. "Weekly Money Announcements: New Information and Its Effects." Federal Reserve Bank of St. Louis *Review* (August/September 1985): 25-34.

Tauchen, George. "Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data." *Journal of Business and Economic Statistics* 4 (October 1986): 397-416.

White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroskedasticity." *Econometrica* 48 (1980): 817-38.



Measuring the Personal Savings Rate: Some Technical Perspectives

R. Mark Rogers

The American personal savings rate has been attracting interest both because of its long-term decline and because of its low level in comparison to that of other industrial countries. However, research into reasons for Americans' savings behavior has not generated robust results. Researchers' difficulties with savings rate explanations may result from inadequate models or incomplete or inaccurate data related to savings behavior, but they may also stem from measurement decisions that on their own can influence the savings rate's observed level and movements over time.

The personal savings rate can differ from one time or place to another for four general reasons: (1) households' behavior may vary; (2) data may be inaccurate or inconsistent in their coverage because of such phenomena as the underground economy; (3) structural economic differences like variations in tax systems may cause savings to shift between the household and corporate sectors; and (4) measurements of elements of personal saving or the savings rate may differ. An extensive group of studies indicates that each of these factors plays a part in intertemporal and intercountry differences.[1] These findings suggest that users of savings rate data should draw conclusions cautiously.

Measurement decisions, the focus of this article, are necessary because records of some components of consumers' income and outlays that are conceptually necessary for measuring the personal savings rate are lacking and because figures collected on other components do not fit the concept that the data on savings are designed to measure. Since records are inadequate, these components must be estimated and appropriately distributed to sectors of the economy. This article, which is not intended to be a comprehensive study of savings rate measurement, is focused on two major types of these components and demonstrates that alternative decisions about whether or how to measure

The author is forecast coordinator in the macropolicy section of the Atlanta Fed's Research Department.

these components have different impacts on both the intertemporal change and level of the personal savings rate. These differences may in turn influence the results of savings behavior analyses.

What is the Savings Rate?

Of published measures of the U.S. personal savings rate, most used is the one computed by the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce.[2] The BEA's broad concept of saving is as a flow of personal income available for but not spent on consumption outlays in the same period in which it is earned. In national income accounting terms, savings equals disposable after-tax personal income less personal outlays.

The BEA does not directly measure savings from actual savings component data series as it does the components of personal income taxes, outlays, and other payments. Using estimated income and outlay components, the BEA calculates the personal savings rate as a percentage of disposable personal income:

$$\text{Personal Savings Rate} = \frac{(\text{Personal Income} - \text{Taxes} - \text{Outlays})}{(\text{Personal Income} - \text{Taxes})} \bullet 100$$

or

$$\text{Personal Savings Rate} = \frac{\text{Personal Savings Rate}}{\text{Disposable Personal Income}} \bullet 100$$

Changes in taxes, outlays, and personal income can thus affect the personal savings rate. Table 1 records the personal savings rate and its components during the 1960-89 period.

In any system of economic data, decisions are made about how various concepts are to be measured in actual practice. Although measures are typically designed to match theoretical concepts as closely as possible, making the data correspond to the concepts is not always clean or precise. Concepts of savings and the savings rate in the BEA's national income accounting system are clear; however, not all of the data are available that would allow a seamless transfer from concept to measure. Since not all flows that make up personal income or all personal outlays during a period are routinely recorded in the private economy, the BEA must devise methods of measuring unrecorded types of income and consumption indirectly with available data. The methods they choose affect the measurement of the residual—savings.

Wage income, for example, is relatively easy to estimate because the U.S. Department of Labor collects earnings data from monthly surveys. Information on proprietors' income is somewhat more difficult to obtain. In both cases, however, data come from reported transactions and reflect the BEA's decisions about what types of data best correspond to its concept of income. Many consumer outlays such as tax payments and purchases at retail stores are also recorded in such a way that direct estimates are straightforward.[3]

Unfortunately, other components of consumer income, taxes, and outlays are not so easily determined. One set of these flows occurs when income is received "in kind"—that is, not as a monetary payment but as a good or service. In-kind flows include food produced and consumed on the farm, meals provided by businesses at little or no cost to employees, and banking services rendered to consumers without an explicit fee. Since such goods and services are in kind, they are consumed simultaneously with the income flow and affect figures for both personal income and outlays—and hence have no effect on the level of personal savings.

Another set of income and outlay flows that is not directly measurable arises from consumers' use of long-lasting products such as owner-occupied housing, autos, and other consumer durable goods, purchased all at once but consumed during a long period. Data reflecting the value of these products are recorded at the time of ownership transfer, but the value of services flowing from a house, an automobile, or a consumer durable during each period is not recorded.

These flows that are not directly measurable may be treated in several ways. In the

Table 1.
Derivation of Personal Savings Rate

(levels in billions of current dollars)

	Personal Income	–	Personal Tax and Nontax Payments	=	Disposable Personal Income	–	Personal Outlays	=	Personal Savings	The Savings Rate: Personal Savings as a Percentage of Disposable Personal Income
1960	409.4		50.5		358.9		338.1		20.8	5.8
1961	426.0		52.3		373.8		348.9		24.9	6.6
1962	453.2		57.0		396.1		370.2		25.9	6.5
1963	476.4		60.5		415.8		391.2		24.6	5.9
1964	510.2		58.8		451.4		419.9		31.6	7.0
1965	552.0		65.2		486.8		452.5		34.3	7.0
1966	600.8		74.9		525.9		489.9		35.9	6.8
1967	644.6		82.5		562.1		516.9		45.2	8.0
1968	707.2		97.7		609.6		567.1		42.4	7.0
1969	773.0		116.3		656.7		614.5		42.2	6.4
1970	831.8		116.2		715.6		657.8		57.8	8.1
1971	894.0		117.3		776.8		710.4		66.3	8.5
1972	981.6		142.0		839.6		778.2		61.3	7.3
1973	1,101.7		151.9		949.8		860.8		89.0	9.4
1974	1,210.1		171.8		1,038.3		941.6		96.7	9.3
1975	1,313.4		170.6		1,142.8		1,038.1		104.6	9.2
1976	1,451.4		198.7		1,252.6		1,156.8		95.8	7.6
1977	1,607.5		228.2		1,379.3		1,288.6		90.7	6.6
1978	1,812.4		261.2		1,551.2		1,441.0		110.2	7.1
1979	2,034.0		304.7		1,729.3		1,611.2		118.1	6.8
1980	2,258.5		340.6		1,918.0		1,781.1		136.9	7.1
1981	2,520.9		393.3		2,127.6		1,968.2		159.4	7.5
1982	2,670.8		409.3		2,261.4		2,107.5		154.0	6.8
1983	2,838.6		410.5		2,428.1		2,297.5		130.6	5.4
1984	3,108.8		440.2		2,668.6		2,504.5		164.1	6.1
1985	3,325.4		486.6		2,838.7		2,713.3		125.4	4.4
1986	3,526.2		512.9		3,013.3		2,888.5		124.9	4.1
1987	3,766.4		571.7		3,194.7		3,102.3		92.5	2.9
1988	4,070.8		591.5		3,479.3		3,333.6		145.7	4.2
1989	4,384.3		658.8		3,725.5		3,553.7		171.8	4.6

Source: Calculated by the Federal Reserve Bank of Atlanta from data from the U.S. Department of Commerce, Bureau of Economic Analysis.

BEA estimates, for instance, income from and consumption of owner-occupied housing are based on an estimate of rent that would have been paid had the housing been rented instead of owned. This procedure results in imputed income and outlay estimates similar to but more complicated than estimates associated with in-kind income and consumption. For autos and other consumer durables, the BEA takes a different tack. It uses current expenditure data as its outlay estimate.

Major Measurement Decisions and Their Effects

This article is a look at the impact of two specific sets of BEA measurement decisions: (1) treatment of imputed income and consumption and (2) treatment of consumption of autos and other consumer durables.

Imputations for Owner-Occupied Housing. In figures for the period 1960-89, imputations of income and outlays related to owner-occupied housing are the largest imputations in personal income and consumption.[4] They, along with imputation of income from financial services received without payment, make up most of the value of personal income imputations. In 1989, imputed gross space rent equaled more than 8 percent of personal income and slightly more than 10 percent of personal outlays.

For both nonfarm and farm components, the imputations for owner-occupied housing begin with estimates of the rental equivalent for housing consumption. The BEA treats homeowners as unincorporated business entities owning residential property and renting it to themselves. Gross earnings are figures for space rent from which estimated costs of maintenance and repair products and services are subtracted to get the contribution of owner-occupied housing to gross housing product. To calculate the net rental income of persons three other items are subtracted—capital consumption allowances, property tax liabilities, and interest payments on housing—and housing subsidies are added. The result, net rental income for owner-occupied housing, is included in personal income. To derive rental expense, imputed rental outlays equaling net rental income plus capital consumption allowances and indirect business taxes are added to personal consumption. Finally, to avoid double counting, the BEA shifts figures for actual purchases of owner-occupied housing out of the personal consumption category. This last change limits the amount of personal outlays on this housing component to imputed rental expense.

As a result of this treatment, personal income figures are reduced by the amount of subsidies; in recent years measured personal income has also been reduced by the amount of net rental income because that number has been negative. Additionally, imputed personal tax payments are less than the nonimputation level because property taxes on owner-occupied housing are charged to the business sector rather than to households. Households' outlays are increased by the amount of imputed rental expenses. The net effect of these imputations is to reduce personal tax payments, thereby raising disposable personal income. Typically in the United States personal consumption outlay figures are reduced because purchases of owner-occupied housing exceed imputed rental expense. Thus, in the U.S. economy measurement decisions have the effect of raising the observed level and rate of personal savings.

Other Imputations. The next largest dollar value of imputed personal income and outlays involves certain financial services received by persons without payment for them. In 1989 this series equaled about 2 percent of personal income and about 2.5 percent of personal outlays. Consumers pay directly for some financial services but not for all. For example, many bank accounts are exempt from service charges if certain minimum balance requirements are met. In addition, a customer who maintains a minimum balance might receive a small insurance policy and additional financial services. The customer's financial institution covers the costs of providing these services with earnings from funds deposited there. In fact, bank interest earnings are typically higher than the combined expenses of these "free services" and interest paid to consumers. Theoretically, a depositor could receive interest income equal to that earned by the bank and could separately purchase the services. The BEA treats the difference between the bank interest earnings on customer deposits and the above two expenses as imputed interest income to persons. An equal dollar amount is imputed to personal consumption for consumption of financial services. For their imputations, the income is equal to the outlays.

Other imputations include employment-related series for food furnished to employees (including military and domestic employees) and standard clothing issued to military personnel. Imputations are also made for rental value of buildings and equipment owned and used by nonprofit institutions serving individuals, for farm products consumed on farms, and for margins on owner-built homes.

While the BEA's imputation decisions are all reasonable, there are alternatives that incorporate other ways of accounting for imputations.[5] The significant point is that the methods of imputation used by the BEA have an impact on both the time path and level of the U.S. savings rate.

Impact of Imputations. What impacts have imputed income and outlays had on the personal savings rate? Table 2 compares levels of personal income and components according to BEA measures, with and without imputations. The inclusion of imputations by the BEA significantly raises total as well as disposable personal income; personal outlays are also increased, although by relatively less. Consequently, as mentioned earlier, personal savings figures are higher with BEA imputations than if imputed income and expenses were ignored.

In their influence on the level of personal and disposable personal income, financial services furnished without payment accounted for virtually all of the higher level. With imputations, personal income was $82.6 billion higher in 1989; of this amount, $90.3 billion came from the financial services component. Owner-occupied housing imputations (farm and nonfarm) reduced personal income by $18.0 billion that year.

Table 2.
Personal Income and Components, Official and Without Imputations
(billions of current dollars)

	Personal Income		Personal Tax and Nontax Payments		Disposable Personal Income		Personal Outlays		Personal Savings	
	Official	Without	Official	Without	Official	Without	Official	Without	Official	Without
1960	409.4	393.2	50.5	56.1	358.9	337.1	338.1	331.9	20.8	5.2
1961	426.0	409.3	52.2	58.5	373.7	350.9	348.9	340.5	24.8	10.3
1962	453.2	435.8	57.0	63.8	396.1	372.0	370.2	361.9	25.9	10.1
1963	476.3	458.1	60.5	68.0	415.8	390.1	391.2	382.8	24.6	7.3
1964	510.2	491.8	58.8	67.0	451.4	424.8	419.9	410.8	31.6	14.0

1965	552.0	532.7	65.2	74.1	486.8	458.5	452.5	441.9	34.3	16.6
1966	600.8	579.7	74.9	84.5	525.9	495.2	489.9	474.6	35.9	20.5
1967	644.6	622.1	82.5	93.2	562.1	528.9	516.9	498.8	45.2	30.2
1968	707.2	684.4	97.7	109.8	609.6	574.6	567.1	548.6	42.4	26.0
1969	773.0	748.9	116.3	129.8	656.7	619.1	614.4	594.1	42.2	25.0
1970	831.8	806.3	116.2	131.6	715.6	674.7	657.8	632.8	57.8	42.0
1971	894.0	868.0	117.2	134.2	776.8	733.8	710.4	693.1	66.3	40.7
1972	981.6	955.3	142.0	160.2	839.6	795.1	778.2	766.3	61.3	28.5
1973	1,101.7	1,073.7	151.9	171.6	949.8	902.0	860.8	848.7	89.0	53.4
1974	1,210.1	1,179.5	171.8	192.8	1,038.3	986.7	941.6	918.1	96.7	68.7
1975	1,313.4	1,280.0	170.6	193.7	1,142.8	1,086.3	1,038.1	1,006.8	104.6	79.5
1976	1,451.4	1,417.0	198.7	224.2	1,252.6	1,192.8	1,156.8	1,137.2	95.8	55.6
1977	1,607.5	1,574.0	228.2	256.2	1,379.3	1,317.8	1,288.6	1,285.1	90.7	32.8
1978	1,812.4	1,771.7	261.2	289.9	1,551.2	1,481.8	1,441.0	1,438.8	110.2	43.0
1979	2,034.0	1,990.6	304.7	334.3	1,729.3	1,656.3	1,611.2	1,589.9	118.1	66.4
1980	2,258.5	2,213.5	340.6	372.6	1,918.0	1,840.9	1,781.1	1,736.8	136.9	104.1
1981	2,520.9	2,475.5	393.3	429.1	2,127.6	2,046.4	1,968.2	1,917.5	159.4	128.9
1982	2,670.8	2,623.3	409.3	448.8	2,261.4	2,174.5	2,107.5	2,037.6	154.0	136.9
1983	2,838.6	2,784.1	410.5	453.0	2,428.1	2,331.1	2,297.5	2,269.9	130.6	61.2
1984	3,108.8	3,059.4	440.2	486.2	2,668.6	2,573.2	2,504.5	2,484.2	164.1	89.0
1985	3,325.4	3,270.0	486.6	536.1	2,838.7	2,733.9	2,713.3	2,687.6	125.4	46.3
1986	3,526.2	3,455.1	512.9	565.6	3,013.4	2,889.5	2,888.5	2,872.0	124.9	17.5
1987	3,766.4	3,679.3	571.7	627.6	3,194.7	3,051.7	3,102.3	3,069.4	92.5	–17.7
1988	4,070.8	3,983.4	591.5	650.6	3,479.3	3,332.9	3,333.6	3,305.8	145.7	27.1
1989	4,384.3	4,301.7	658.8	723.1	3,725.5	3,578.6	3,553.7	3,507.2	171.8	69.4

Source: Calculated by the Federal Reserve Bank of Atlanta from data in *Survey of Current Business*, Table 8.9, July issues of 1986, 1987, 1988, 1989, and 1990. *National Income and Product Accounts of the United States, 1929-82,* Table 8.9.

Outlays were also higher with imputations. Of course, the treatment of financial services received without payment raises both income and consumption by the same amount, giving a greater relative boost to outlays. However, a significant portion of personal outlays is credited toward net purchases of housing. This calculation reduces the overall amount by which the imputations for outlays raise the level of personal outlays. In 1989, imputations raised outlays by only $46.5 billion, compared with an $82.6 billion increase in income.

Finally, personal tax and nontax payments are considerably lower with imputations included. The difference was $64.3 billion in 1989. The factor accounting for the lower figures is the shifting of taxes and interest expenses to owner-occupied housing charges against GNP.

The personal savings rate is also higher (see Table 3). The BEA's mean rate for the 1960-89 period is 6.7 percent. Without imputations the mean is 3.8 percent. The savings rate's small downward trend over the entire 1960-89 period is about one and three-quarters times the without-imputations trend. Imputations also considerably reduced the savings rate's variation. The official series' coefficient of variation is less than half that of the series without imputations. These impacts on level, trend, and variation can also be seen in Chart 1. Interestingly enough, both series declined by 6.5 percentage points from the post-1960 peak savings rate in 1973 to the 1987 low point.

Treatment of Autos and Other Consumer Durables

Those who monitor monthly releases of economic data have long recognized that swings in sales of autos and other consumer durables have a significant impact on changes in measurements of personal consumption and an even greater relative effect on the personal savings rate. The BEA's treatment of consumption outlays for autos and other consumer durables accounts for this pattern: the BEA counts auto and other durables purchases by consumers as fully consumed in the month of purchase. This approach simplifies the measurement of auto and consumer durables production as part of the BEA's GNP estimate, which is a measure of current production, not income. Other measurement decisions would yield different time patterns of income and consumption and different levels of the savings rate.

An alternative approach to that used by the BEA is to estimate the flow of services from autos and other consumer durables consumed over time following the principle behind the imputation of income from and consumption of the services of owner-occupied housing.[6] Using an estimate of depreciation to measure the flow of income from the consumption of consumer durables in years beyond the purchase date illustrates this alternative's impact on savings rate measures.[7]

To determine the method's relative influence on the personal savings rate, an adjusted personal savings rate series can be created. Using a portion of current consumer durables consumption and adding a share of sales from prior years over the depreciable lifetime of the product, one can calculate approximately the impact of measuring income and services consumed from consumer durables.

For automobiles five years is the appropriate depreciation period; personal consumption for motor vehicles and parts is estimated on a straight-line basis. Assuming that purchases on average take place in mid-year (that is, purchases are evenly distributed over the year), one-tenth of consumption takes place in the year of purchase, another one-fifth occurs in each of the next four years, and the remaining one-tenth is depreciated the sixth year. To create the adjusted personal savings rate series, the official component for motor vehicle consumption is subtracted from outlays and the new estimate using this depreciation measure is substituted. Table 4 and Chart 2 compare the personal savings rate measured with the adjustment for auto income and consumption.

As shown in Table 4, adjusted in this manner, the series for the personal savings rate

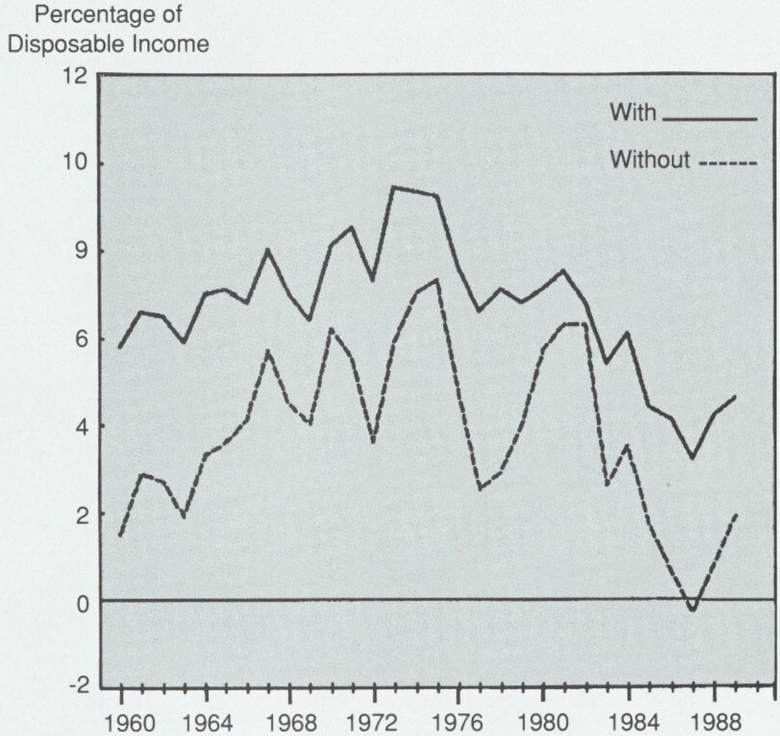
Table 3.
Personal Savings Rate, Official and Without Imputations

	Personal Savings Rate			Year-to-Year Change		
	Official	Without	Difference	Official	Without	Difference
1960	5.8	1.5	4.3	–0.5	0.1	–0.6
1961	6.6	2.9	3.7	0.8	1.4	–0.6
1962	6.5	2.7	3.8	–0.1	–0.2	0.1
1963	5.9	1.9	4.0	–0.6	–0.8	0.2
1964	7.0	3.3	3.7	1.1	1.4	–0.3
1965	7.1	3.6	3.5	0.1	0.3	–0.2
1966	6.8	4.1	2.7	–0.3	0.5	–0.8
1967	8.0	5.7	2.3	1.2	1.6	–0.4
1968	7.0	4.5	2.5	–1.0	–1.2	0.2
1969	6.4	4.0	2.4	–0.6	–0.5	–0.1
1970	8.1	6.2	1.9	1.7	2.2	–0.5
1971	8.5	5.5	3.0	0.4	–0.7	1.1
1972	7.3	3.6	3.7	–1.2	–1.9	0.7
1973	9.4	5.9	3.5	2.1	2.3	–0.2
1974	9.3	7.0	2.3	–0.1	1.1	–1.2
1975	9.2	7.3	1.9	–0.1	0.3	–0.4
1976	7.6	4.7	2.9	–1.6	–2.6	1.0
1977	6.6	2.5	4.1	–1.0	–2.2	1.2
1978	7.1	2.9	4.2	0.5	0.4	0.1
1979	6.8	4.0	2.8	–0.3	1.1	–1.4
1980	7.1	5.7	1.4	0.3	1.7	–1.4
1981	7.5	6.3	1.2	0.4	0.6	–0.2
1982	6.8	6.3	0.5	–0.7	0.0	–0.7
1983	5.4	2.6	2.8	–1.4	–3.7	2.3
1984	6.1	3.5	2.6	0.7	0.9	–0.2
1985	4.4	1.7	2.7	–1.7	–1.8	0.1
1986	4.1	0.6	3.5	–0.3	–1.1	0.8
1987	2.9	–0.6	3.5	–1.2	–1.2	0.0
1988	4.2	0.8	3.4	1.3	1.4	–0.1
1989	4.6	1.9	2.7	0.4	1.1	–0.7
Mean	6.68	3.75				
Coefficient of Variation	0.23	0.54				
Trend	–0.08	–0.05				
Long-Term Decline 1987–73:	6.5	6.5				

Note: The coefficient of variation is the ratio of the standard deviation to the mean. Trend is the coefficient of a time variable in an ordinary least squares regression with the personal savings rate, the dependent variable and the independent variables being time and a constant. Time equals 1 for 1960, 2 for 1961, and 3 for 1962, and so forth.

Source: Calculated by the Federal Reserve Bank of Atlanta from data from the U.S. Department of Commerce, Bureau of Economic Analysis.

Chart 1.
Personal Savings Rates, With and Without Imputations
1960-89



Source: See Table 3.

has a higher value throughout than does the official data series. The coefficients of variation are slightly lower than the official rate, and the decline from 1973 through 1987 is slightly higher—6.8 percentage points for the adjusted series versus 6.5 percentage points for the unadjusted data. Apparently, the BEA's treatment of auto purchases was somewhat favorable in terms of the decline in the personal savings rate's not being as severe over the 1973-87 period with the official series versus the adjusted series. However, the adjustment significantly increased the estimated savings rate by an average of 1 percentage point over the 1960-89 period.

Table 4 also shows that applying the procedure used with motor vehicles to estimate an adjusted depreciation series for autos and other durables significantly raises the personal savings rate above official and auto-adjusted levels. The nonauto adjustment provided a little more than half of the upward boost of all durables. Variability also is marginally less. The personal savings rate decline also becomes more pronounced when the savings rate is adjusted for autos and other durables. Thus, the BEA's procedure of counting durables as fully consumed when purchased was slightly favorable toward the observed personal savings rate during the 1973-87 period. On the other hand, the use of durables consumption based on depreciation raises the personal saving rate. This series yielded a figure 2.9 percentage points higher than the BEA's in 1973 and 2.3 percentage points in 1987.

The above procedure is a simple method of adjusting for durables consumption and investment. In fact a more elaborate method suggested by Blades and Sturm

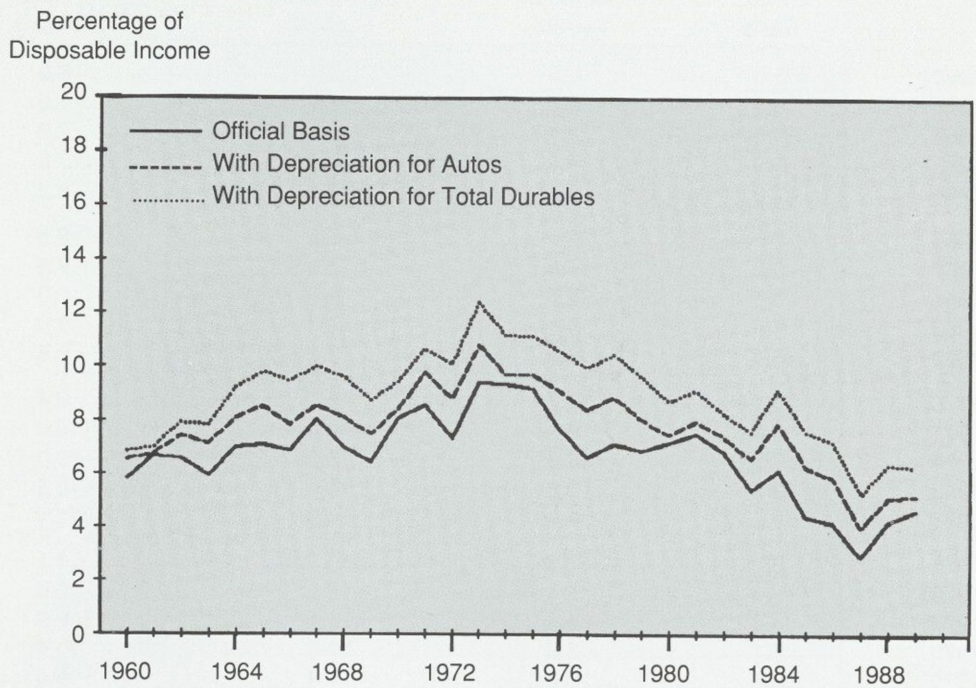
Table 4.
Personal Savings Rates, With Depreciated Durables Consumption

	Official Basis	With Depreciation for Motor Vehicles	With Depreciation for Nonauto Durables	With Depreciation for Total Durables
1960	5.8	6.5	6.1	6.8
1961	6.6	6.7	6.9	7.0
1962	6.5	7.4	7.0	7.9
1963	5.9	7.1	6.6	7.8
1964	7.0	8.1	8.1	9.2
1965	7.0	8.5	8.3	9.8
1966	6.8	7.8	8.5	9.4
1967	8.0	8.5	9.5	10.0
1968	7.0	8.1	8.4	9.6
1969	6.4	7.5	7.7	8.8
1970	8.1	8.4	9.1	9.4
1971	8.5	9.8	9.4	10.6
1972	7.3	8.8	8.6	10.0
1973	9.4	10.8	10.9	12.3
1974	9.3	9.7	10.8	11.1
1975	9.2	9.7	10.5	11.1
1976	7.6	9.1	9.1	10.6
1977	6.6	8.4	8.1	9.9
1978	7.1	8.8	8.7	10.4
1979	6.8	8.0	8.4	9.6
1980	7.1	7.4	8.4	8.7
1981	7.5	7.9	8.6	9.1
1982	6.8	7.4	7.7	8.2
1983	5.4	6.6	6.4	7.6
1984	6.1	7.9	7.4	9.1
1985	4.4	6.3	5.7	7.7
1986	4.1	5.9	5.5	7.2
1987	2.9	4.0	4.2	5.2
1988	4.2	5.1	5.4	6.4
1989	4.6	5.2	5.7	6.3
Mean	6.7	7.7	7.9	8.9
Coefficient of Variation:	0.23	0.19	0.21	0.19
Trend	–0.08	–0.07	–0.06	–0.05
Long-Term Decline 1987-73	6.5	6.8	6.7	7.1

Note: *The coefficient of variation is the ratio of the standard deviation to the mean. Trend is the coefficient of a time variable in an ordinary least squares regression with the personal savings rate, the dependent variable and the independent variables being time and a constant. Time equals 1 for 1960, 2 for 1961, and 3 for 1962, and so forth.*

Source: Calculated by the Federal Reserve Bank of Atlanta from data from the U.S. Department of Commerce, Bureau of Economic Analysis.

Chart 2.
Personal Savings Rates, With Depreciated Durable Consumption 1960-89



Source: See Table 3.

(1982) produces a dramatically higher personal savings rate over the 1960-89 period. Their method essentially adds durables purchases to the official savings level for the numerator of the personal savings rate and adds depreciation of durables purchases as an income component to disposable personal income for the denominator. Using the Blades and Sturm method and the depreciation estimates given earlier, this type of adjustment for total durables produces a mean value of 17.42 percent for the personal savings rate over the 1960-89 period with a trend of -0.06 and coefficient of variation of only 0.08. However, the 1973-87 decline is lessened (in contrast to the other adjustment method), with only a 5.8 percentage point drop. Obviously, the method of adjustment does significantly affect the level of the personal savings rate as well as the long-term trend.

Impacts on Intertemporal Comparisons of the U.S. Savings Rate

The two measurement conventions discussed in this article—the National Income and Product Account treatments of imputed income and consumption and of auto and consumer durables consumption—certainly are not the only ones that affect the personal savings rate. However, the data presented here indicate that measurement decisions related to these components have influenced the level as well as intertemporal variation of the measured savings rate.

The BEA's housing treatment raised the personal savings rate because property taxes as well as expenses were shifted to the business sector and because outlays were reduced by net purchases of housing. Its treatment of autos and consumer durables led to a savings rate lower than the level it would have reached had these goods been treated as providers of consumer services over several years.

Conclusion

This analysis seeks to demonstrate that decisions about the measurement of personal savings rate components can influence levels, variability, and trends in the personal savings rate. Given the influence of some decisions on difficult measurement issues in determining U.S. savings rate trends and levels, the "official" personal savings rate should not be considered an exclusive or a pure reflection of consumer behavior. The usefulness of the personal savings rate measure for analyzing personal savings behavior both over time and among countries can be improved by careful analysis of the impact of measurement decisions on the savings rate's level and trends.

Notes

[1] The following studies exemplify discussions of how various factors influence the personal savings rate. Hypotheses about savings behavior over time in the United States, particularly personal savings, are considered in Summers and Carrol (1987) and by Blinder and others in the discussion accompanying that article. Intercountry differences in behavior are the focus of Carrol and Summers (1987) and Frietas (1981); Blades and Sturm (1982) discuss both economic structure and definition differences, while Hendershott and Peek (1987), deLeeuw (1984), and Ruggles and Ruggles (1986) deal with the influences of varying definitions.

[2] The Board of Governors of the Federal Reserve System (1980) publishes a different measure of the personal savings rate, based on flow of funds data. See deLeeuw (1984) for a discussion of differences between the BEA and Federal Reserve measures and an evaluation.

[3] For more detailed information on how the BEA estimates personal income, see Byrnes et al. (1979).

[4] Data come from U.S. Department of Commerce, *Survey of Current Business*, July 1986-90, Table 8.9, and *National Income and Product Accounts of the United States*, 1929-82.

[5] See, for example, Ruggles and Ruggles (1986) for another approach to imputations of personal income taxes and expenditures.

[6] This approach has been suggested by several other students of national income accounting. For instance, see Blades and Sturm (1982), Ruggles and Ruggles (1986), and Hendershott and Peek (1987).

[7] If treated like purchases of owner-occupied housing, such an estimate would also account for personal property taxes paid on autos and other durables and for repairs, maintenance, and interest expenses. Estimates in the remainder of this paper follow Corrado and Steindel (1980) in ignoring these items. Ruggles and Ruggles (1986) adopt the more complicated method more closely paralleling treatment of owner-occupied housing. Hendershott and Peek (1987) adjust by adding auto and consumer durables outlays to savings.

References

Blades, Derek W., and Peter H. Sturm. "The Concept and Measurement of Savings: The United States and Other Industrialized Countries." In *Saving and Governmental Policy*, Conference Series 25, 1-30. Boston, Mass.: Federal Reserve Bank of Boston, October 1982.

Board of Governors of the Federal Reserve System. *Introduction to the Flow of Funds*. Washington, D.C., 1980.

Byrnes, James C., Gerald F. Donahue, Mary W. Hook, and Robert P. Parker. "Monthly Estimates of Personal Income, Taxes and Outlays." *Survey of Current Business* 59 (November 1979): 18-36.

Carrol, Chris, and Lawrence H. Summers. "Why Have Private Savings Rates in the United States and Canada Diverged?" *Journal of Monetary Economics* 20 (September 1987): 249-79.

Corrado, Carol, and Charles Steindel. "Perspectives on Personal Saving." *Federal Reserve Bulletin* 66 (August 1980): 613-25.

deLeeuw, Frank. "Conflicting Measures of Private Saving." *Survey of Current Business* 64 (November 1984): 17-23.

Freitas, Lewis P. "Comparing the Economic Performance of Japan and the United States." *Journal of Accounting, Auditing, and Finance* 4 (Spring 1981): 270-74.

Hendershott, Patric H., and Joe Peek. "Private Saving in the United States: 1950-85." National Bureau of Economic Research Working Paper 2294, June 1987.

Ruggles, Richard, and Nancy D. Ruggles. "The Integration of Macro and Micro Data for the Household Sector." *Review of Income and Wealth* 32 (September 1986): 245-76.

Summers, Lawrence H., and Chris Carrol. "Why is U.S. National Saving So Low?" *Brookings Papers on Economic Activity* 2 (1987): 607-42.

Book Review

High Tech, Low Tech, No Tech: Recent Industrial and Occupational Change in the South

by William W. Falk and Thomas A. Lyson.
Albany, N.Y.: State University of New York Press, 1988.
181 pages. $44.50.

The South's economic development, or lack thereof, has been the subject of many treatises since economic progress in the region was first recognized as lagging seriously behind advances in the rest of the United States. A number of factors are blamed for contributing to the South's persistent economic backwardness. Most have to do with the region's generally poor use of resources, the low skill levels and earning power of southerners, and the ineffectiveness of measures adopted to remedy the region's deficiencies.

The 1970s brought the first cause for optimism in the South as the miraculous ingredients of robust economic growth seemed suddenly to have come together in the region. The "Sun Belt" became the place of choice for increasing numbers of new residents. Indeed, long-standing problems appeared to be fading, and a new era of faster-than-national economic growth had finally arrived.

This new picture of the South is a mirage, however, according to William W. Falk and Thomas A. Lyson. In *High Tech, Low Tech, No Tech: Recent Industrial and Occupational Change in the South*, the authors argue that the false optimism is based on improved performance of only narrow sectors of the region. They contend that most of the South remains mired in its traditional backwaters that have only been further muddied by misguided efforts to generate economic progress. The South's vast rural areas and their poorly-skilled residents have largely been left behind by the glistening development that has in fact been limited to a few major metropolitan areas, according to the authors.

Falk, who chairs the Department of Sociology at the University of Maryland, College Park, and Lyson, an associate professor of urban sociology at Cornell University, focus on changes since the 1970s to demonstrate that the widely publicized Sun Belt growth has skipped over broad geographic regions and population bands. To demonstrate the disparities of economic growth, the authors separate southern counties into four groups based on the relative concentrations of their rural and urban populations and the relative racial composition of rural areas. One group is composed of counties included in metropolitan statistical areas, labeled MSA counties. A second group consists of non-MSA counties with 20,000 or more urban residents, which are called Other Urban counties. A third group includes the rural counties with mostly white residents, referred to as Rural White counties. Rural counties with relatively large black populations (40 percent or more of total residents) make up what the authors have called the Black Belt counties, the fourth group.

Falk and Lyson's comparisons of industrial activity and occupational distributions among the four county groups provide the primary evidence of economic growth disparities between particular areas and population groups across the South. For example, the two urban county groups have higher proportions of service and trade industries, and the two rural groups exhibit larger shares of farming and manufacturing. While blue-collar jobs are main occupations of rural county residents, professional, managerial, and sales/clerical jobs are more prevalent occupations in the urban counties, especially the MSA group.

The authors claim that these industrial and occupational differences reflect the operations of de facto industrialization policies in the South during the past fifty years. Under the policies, various incentive packages have been offered primarily by state and local governments to attract new business growth. Most of these incentives are said to favor urban development in that they have helped cities and their suburbs attract rapidly growing, technologically advanced industries that provide a large number of good jobs.

Rural areas, on the other hand, while also experiencing growth since 1970, have tended to gain the less dynamic and innovative businesses looking for cheap labor and relocation subsidies such as tax breaks, training programs, and low-cost financing schemes. In some cases, Falk and Lyson contend, the costs to the communities of these incentives have outweighed the advantages of having the new employer move in. Furthermore, the authors note, such firms will frequently, with little warning, move to another location before a community has had time to recover recruitment costs. In spite of this pattern, the philosophy that "any job is better than no job" keeps rural developers on the chase for such companies.

The jobs offered by the highly mobile firms that rural areas attract are typically concentrated in textile, apparel, and other labor-intensive occupations. They require relatively little training, have few benefits, are largely nonunionized, and offer almost no income growth potential. In addition, Falk and Lyson observe, these jobs are vulnerable to being exported to Third World countries where wages are still cheaper and workers even more abundant than in southern rural areas. Consequently, the authors project that industries will continue to move in and out of the rural South in revolving-door fashion unless policies are adopted to make immigrant companies fiscally and ethically responsible.

Another employment problem in the South is revealed in comparisons of the proportions of white and black men and women found in various occupations. Falk and Lyson cite evidence of rampant racial and sexual job discrimination. When the authors divided jobs into desirable and undesirable categories based loosely on the menial nature of the tasks, the variations in jobs held by different race/sex groups were shown to be quite large. The authors found that in general white men were more than two times as prevalent in the better jobs than white women or black men and over four times more than black women. Furthermore, by Falk and Lyson's measures, the more rural the geography, the more exaggerated these disparities became.

Although they provide no evidence to indicate how the South differs from the nation overall or from other regions, the authors attribute employment inequities among race/sex groups to the entrenchment of traditional southern socioeconomic patterns. They conclude that the industrial recruitment strategies used to attract firms from outside the region have done surprisingly little to alter the social landscape and parallel economic picture of the South.

What is to be done about the problems that hinder the South's economic progress, especially in the predominant rural areas? The authors suggest two new policy thrusts for future development. They call for a national industrial policy to coordinate and complement activities at the national, state, and local levels, striving for an equitable distribution of jobs across rural and urban sectors of a region. Falk and Lyson envision this policy as one that would encourage the existing industrial base to remain and grow, helping new firms get started and recruiting high-growth industries that would match and upgrade the skills of the local labor market. Such an approach would also involve abandoning the short-sighted and ultimately self-defeating practices of offering ever-more-costly incentives in an effort to capture footloose firms.

The authors recommend other initiatives, including some suggested by bodies such as the Southern Growth Policies Board, that would concentrate on enhancing human resource development through improved education. Falk and Lyson recommend a restructuring of vocational educational programs in rural communities from specific job training toward development of more general skills and behaviors affecting work performance—independence, responsibility, entrepreneurship, and the like. Specific job preparation would be shifted to on-the-job training in the private sector. Additionally, Falk and Lyson recommend development of small business incubator facilities to nurture the creation and expansion of new and small businesses in rural areas. Special mechanisms would channel venture or risk capital to these enterprises. Finally, the authors suggest that a national industrial policy address the obstacles faced by various disadvantaged worker segments and provide for creation of more desirable jobs in rural areas.

Falk and Lyson's final policy suggestion is to adopt what they call democratic economic planning. By this they mean abandoning current market-based allocation of goods and services and relying instead on a system of direct planning that focuses on meeting the populace's needs for goods and for personal development. A major component of this system would be the development of public enterprises at all government levels to provide the goods and services that private enterprises cannot or will not produce.

The authors' ideas of democratic economic planning include programs that would preserve industries and businesses whose economic livelihood is threatened because their products can be produced more cheaply or of better quality elsewhere. These programs would help employees purchase floundering plants. Falk and Lyson contend that, under employee ownership and management, such plants could be operated more profitably and preserved, thereby saving jobs and at the same time giving employees more control over their economic destinies.

Falk and Lyson close their treatise on southern economic problems with a recitation of the region's more recent economic afflictions during the 1980s—the drop in oil

"[T]he authors suggest that a national industrial policy address the obstacles faced by various disadvantaged worker segments and provide for creation of more desirable jobs in rural areas."

prices, the closing of textile mills, the loss of southern lumber markets to subsidized Canadian lumber, the farm crisis attributable to both drought and loss of overseas markets, and the increasing movement of branch plants from the South to Third World locations. The authors argue that all of these developments further signify the need for greater equality of opportunity for economic success to be afforded to everyone. To achieve this goal, they assert, the way American business is conducted and federally regulated must be radically altered. Falk and Lyson condemn the operation of American-style capitalism for its exclusive profit motive, which they maintain has less concern for the common good than for the welfare of a few (presumably the business owners), thereby foreclosing the possibility of an economically just society.

While Falk and Lyson have identified some of the South's important social issues and serious obstacles to economic development, one must question several of the authors' assertions and data. Are the differences observed among population groups in geographical areas in fact attributable solely to current racial, sexual, and geographical discrimination, or may they be, at least partially, the legacy of past patterns of discrimination in jobs and education? Conspicuous by its absence is any reference to educational differences among the various population groups. Although the authors acknowledge the need for higher levels of education as a prerequisite to economic progress, they fail to compare the educational attributes of their urban, rural white, and rural black groups. It is reasonable to expect greatest economic development in areas where educational levels of the populace are highest. It is also logical that better jobs would be held by those individuals best prepared to perform those jobs, yet Falk and Lyson do not adequately consider current variations in education among population groups.

"Although the authors acknowledge the need for higher levels of education as a prerequisite to economic progress, they fail to compare the educational attributes of their urban, rural white, and rural black groups."

For example, in Georgia, whose mix is reasonably representative of the South as a whole, there are interesting corollaries between educational attainment of the adult population in various county regions and the economic opportunities in those areas as cited by Falk and Lyson. By this reviewer's analysis, the highest average percentage of people having eight or fewer years of schooling (37.7 percent) was found in Georgia's thirty-five nonmetropolitan counties where black residents make up 40 percent or more of the total population. These counties coincide with the authors' rural black counties. Comparable percentages in their eighty-six nonurban white counties and the thirty-eight urban counties are 33.7 and 23.2, respectively.

Proportions of high school graduates in Georgia's population were 22 percent in rural black counties, 26.1 percent in nonurban white counties, and 30.5 percent in metropolitan counties. Thus, the highest proportions of people with fewer years of education are in rural areas of the state, and the heaviest concentrations are in the rural black areas. It is only in the category of four years or more of college that rural black counties approach the rural white counties, but the proportions (7.7 percent versus 8.5 percent) are so low as to be of minor consequence compared to the 14.6 percent with four years or more of college education in metropolitan areas. Educational quality may actually be lower in rural school districts with more limited resources for education. Thus, one would expect low-skill employment opportunities to be the most prevalent job types in the rural areas.

Falk and Lyson also demonstrate that women hold fewer well-paying, more desirable jobs than men in all areas, a fact they attribute entirely to current sexual discrimination. However, educational differences may be an explanatory factor here as well. In Georgia, a smaller percentage of women than men twenty-five years old and over have completed high school and college. The difference is greatest at the college level; 22.8 percent of the men versus 16.0 percent of women had completed college in 1987. Also, in Florida and North Carolina, other southern states for which data are available, significantly higher proportions of men than women had completed college.

The educational differences between men and women were apparently not the same for the black subgroup of the population, however. At the national level, black women and men were more nearly equal in educational attainment, with women holding a slight edge in high school and early college education. Assuming that relationship also holds for the South, the fact that black women occupy the lowest rung on the job desirability ladder as described by Falk and Lyson is consistent with a pattern of sexual discrimination. However, it is also possible that the educational pursuits of black women have not qualified them suitably for such skilled occupations as those of plumbers, electricians, welders, machinery operators, and machinists, which are the higher-paying jobs offered by the manufacturers in rural areas where women are most disadvantaged.

In assessing the impact of current discrimination, one is left to wonder why manufacturers from outside the region (and presumably free of the South's traditional biases) would do other than fill their jobs, especially those in the higher wage categories, with the people available who are likely to be most productive. Falk and Lyson might more accurately have reflected current discrimination had they presented evidence of variation in wage rates or pay for a particular job by sex or racial group. Unfortunately, detailed data are not generally available to examine wage variability by sex and race within particular occupational groups. A few other researchers, however, have explored closely related topics.

A recent study by James J. Heckman and Brook S. Payner used a unique source of data on industrial employment in South Carolina to study employment and wages by race, sex, and industry.[1] The researchers found that real wages differed among race/sex groups in a pattern rather consistent with the "good" job–"bad" job employment distribution of these groups as presented by Falk and Lyson. That is, from 1940 to 1967, real wages were highest for white men and next highest for white women; black men ranked third, and black women were at the bottom of the scale. Heckman and Payner found that manufacturing employment of blacks remained at relatively low levels until the mid-1960s. However, they found that once skill levels were accounted for, blacks were not underrepresented in economic sectors other than textiles even in 1960. They also found evidence that as their skill levels expanded, so did blacks' employment and average wage rates.

Falk and Lyson's ignoring evidence that current education levels are related to employment distribution and wages probably places too much emphasis on discrimination as an explanation of current race and sex economic differences. Different education and skill levels seem to play a part, although past discrimination may have been quite influential in education and skill attainment. By their emphasis on discrimination the authors actually undercut their recommendation that southern governments use more and better resources to provide education as an important contribution to southern economic development.

Perhaps the most disturbing portions of the Falk and Lyson book from an economic viewpoint are those which communicate the authors' flagrant disregard for the role of differential resource endowments in economic development and the proposal to abandon the marketplace as a director for the employment of those resources. Sustainable economic development of an area must be based on advantages that place it in the most favorable (or least unfavorable) position with competing locations. In the past, an abundance of low-skilled labor has been the principal comparative advantage of many southern areas in attracting labor-intensive industries. Plentiful labor meant low labor costs for firms such as textile and apparel manufacturers for whom low-skilled labor represented a high proportion of production costs. Such firms moved away from higher labor cost areas in the Northeast and into the South, where workers often changed willingly from farms, where their incomes were even less, to jobs at these firms. The low market rate for labor was the signal that directed the inflow of labor-intensive industries, not the decree of a planning group that determined that textile and apparel firms should move to the South to produce a more balanced economy.

The market system also works to draw mobile resources away from areas where their abundance assigns them relatively low market value and into areas where they are relatively more scarce and have greater value. Higher wage rates in urban versus rural areas and in the industrial North versus the rural South have been responsible for massive waves of human migration away from low-wage localities to areas with more lucrative employment opportunities. It seems clear that poverty, especially in southern rural areas, would be much more severe than it is now if the populace had been prevented from responding to market forces that drew workers to better opportunities.

The free market system is not a panacea, of course. Public programs to encourage the transfer of workers to other locations, to retrain or educate workers for new occupations, and to supplement incomes of people who cannot become employable provide direct benefits to workers and to society. The market system alone often will not provide this assistance. The most desirable public sector programs work to upgrade people's chances of gaining employment so that they need rely only temporarily on public assistance. By contrast, programs that encourage inefficient workers to remain in noncompetitive and unprofitable industries, like some Falk and Lyson describe, for indefinite periods are likely to work to the long-term detriment of the very people, industries, and areas intended for assistance. Rather than being a solution, such subsidies tend to discourage the adjustments needed to regain economic efficiency. For that reason, Falk and Lyson's proposal to abandon market-based allocation of goods and services in favor of a direct planning system could be dangerously counterproductive.

The future development of the South is likely to be based on advantages other than cheap labor since even the lowest-paid southern workers are now much more costly than laborers in Third World countries. Such attributes as low-cost land and facilities, abundant fresh water, easy transportation access, minimal congestion, and a less stringent regulatory environment may be the drawing cards that increasingly account for new business locations in rural areas. It will be some combination of such factors that reduces cost, enhances output, or increases monetary and aesthetic returns to entrepreneurs and attracts new business to an area. In the existing governmental (and economic) system, no amount of the authors' prescription for "democratic economic planning" is likely to override the influence of basic economic forces on economic development.

Gene D. Sullivan

The reviewer is a research officer in the regional section of the Atlanta Fed's research department.

Note

[1] James J. Heckman and Brook S. Payner, "Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina," *The American Economic Review* 79 (March 1989): 138-77.

CRA Brochure Details New Ratings, Availability Provisions

To help consumers better understand the Community Reinvestment Act and recent changes in that legislation requiring banks to meet the credit needs of their entire communities, the Federal Reserve Bank of Atlanta has published *Is My Bank Meeting Its Community Reinvestment Obligations?*

The pamphlet outlines the CRA, describes what it requires of banks and thrifts, and explains changes that took effect July 1, 1990, making each bank's most recent rating available. Directed toward the general public, the publication is designed to help consumers comprehend how regulators arrive at an overall CRA rating for an institution. The brochure may also be used as a vehicle for banks to communicate recent CRA modifications to their communities and customers.

Up to 50 copies may be ordered free. Additional copies are available for $.15 each (educators exempt). To order, write to:

Public Information Department
Federal Reserve Bank of Atlanta
104 Marietta Street, N.W.
Atlanta, Georgia 30303-2713
404/521-8788.

Economic Review

Federal Reserve Bank of Atlanta
104 Marietta St., N.W.
Atlanta, Georgia 30303-2713

Address Correction Requested

ALDR LB 2
CAROL ALDRIDGE
RESEARCH LIBRARY
FED RES BANK OF PHILADELPHIA
10 INDEPENDENCE
PHILADELPHIA PA 19106

Bulk Rate
U.S. Postage
PAID
Atlanta, Ga.
Permit 292