CAN YOU DERIVE MARKET VOLATILITY FORECASTS FROM THE OBSERVED YIELD CURVE CONVEXITY BIAS?

WESLEY PHOA

The yield on a thirty-year Treasury bond is often, but not always, lower than the yield on a twenty-year bond. A common and plausible explanation is that since the thirty-year bond has more convexity than the twenty-year bond, it will tend to outperform when bond yields fluctuate. Thus investors pay for convexity by accepting a lower yield.

The convexity bias in observed yields has been the subject of much attention in the swap market; see Burghardt and Hoskins [1995] for a popular exposition. In this case, it is observed that the ten-year swap rate is significantly lower than the yield on an equivalent ten-year strip of Eurodollar futures. This is because the swap has convexity, while the futures position does not; a dealer who is receiving swap and short Eurodollar futures benefits from being long convexity.

There are various rules of thumb for computing an appropriate yield adjustment to take this convexity effect into account; I describe a simple but robust formula later. One could also take a more rigorous approach, and derive the theoretical convexity bias using a formal term structure model.

Since convexity delivers benefits only if yields are volatile, the size of the convexity adjustment depends on how volatile yields are expected to be. There are thus two approaches one can take. First, one can estimate volatility separately, and then use this to calculate what the appropriate convexity adjustment should be.

Alternatively, one can observe actual swap or bond yields, and attempt to deduce what level of volatility is implied by those yields, i.e., market participants' expectations about future (medium- to long-term) volatility, as reflected in observed yields. We refer

WESLEY PHOA is vice president of research at Capital Management Sciences in Los Angeles.

to such a volatility estimate as the *implicit volatility*.

Implicit volatility is clearly different from historical volatility. The market's expectations about future volatility are only partly determined by historical experience. Yet since changes in long-term volatility — as opposed to short-term volatility spikes — tend to be triggered by structural shifts in the economy, a large discrepancy between historical volatility and implicit volatility should be regarded as surprising, and worthy of further investigation.

Implicit volatility may also be different from the implied volatilities observed in the interest rate derivatives markets. The value of the convexity in a long bond is fully realized only over a long period of time — ten years or more. (Note that it is irrelevant that the bond may change hands over this period, provided the potential future value of convexity is priced into each transaction.) Short-dated implied volatilities of bond futures options are clearly irrelevant; cap/floor volatilities are not directly relevant, because they are the implied volatilities of money market rates and not bond yields. A comparable implied volatility would be that on (say) a ten-year swaption on a ten-year swap, but prices on such long-dated contracts are not routinely quoted.

Furthermore, because the importance of the convexity bias has been appreciated only relatively recently, there may well be a discrepancy between implicit volatility and OTC option-implied volatilities even where they are directly comparable, such as in the swaps market. In theory, this gives rise to an arbitrage opportunity; in practice, it may be difficult to exploit a convexity arbitrage efficiently.

This study focuses on bond market data and attempts to estimate implicit volatility from long bond yields. For reasons explained in section I — briefly, the need to incorporate the effect of the overall yield curve slope, to discount the effect of yield anomalies on specific bonds, and to take into account the sometimes aberrant behavior of the short end of the yield curve — the method adopted may appear somewhat roundabout.

Section I describes a method of parameterizing yield curve data based on a formal, economically justified theory of the structure of investor expectations about future bond market returns; the parameters are short- and long-run expected returns, a rate of adjustment coefficient, and the implicit volatility that we wish to estimate. I then describe the estimation procedure, a constrained non-linear optimization problem that turns out to require some care, and present the results.

I relate the results to historically observed bond yield volatilities and describe some possible interpretations, while the conclusion outlines some possible implications for long bond trading strategies with a medium-term horizon as well as some future areas of research.

I have tried to use methods that are as simple and direct as possible. While there is some subtlety in the formulation of concepts, the technical details are mostly straightforward. The exposition should be readily accessible to investment managers who may be unfamiliar with the mathematics of term structure models.

I. FORMALIZATION OF THE EXPECTATIONS THEORY

This study makes use of constant-maturity Treasury yields provided by the Federal Reserve Board (see Exhibit 1). It is thus possible to study a relatively long history, going back to 1953 and extending over several radically different periods in the development of the U.S. fixed-income markets. Note that although some yield histories are available back to 1921, long bond yields for specific maturities are not provided, making it impossible to extend the study beyond 1953 using these data.

Although the convexity bias in long bond yields is most apparent in the twenty- to thirty-year yield spread, there are a number of reasons why it is undesirable to adopt a method of analysis that focuses solely on this spread. The most obvious reason is that this would restrict us to studying periods when both twenty-year and thirty-year historical yields were available, severely

EXHIBIT 1 ■ U.S. Treasury Yield Data Used

	Monthly Series		Daily Series	
	From	To	From	To
1 year	Apr–53	Sep–96	2/15/1977	9/30/96
2 year	Jun–76	Sep–96	2/15/1977	9/30/96
3 year	Apr–53	Sep–96	2/15/1977	9/30/96
5 year	Apr–53	Sep–96	2/15/1977	9/30/96
7 year	Jul–69	Sep–96	2/15/1977	9/30/96
10 year	Apr–53	Sep–96	2/15/1977	9/30/96
20 year (old)	Apr–53	Dec–86	—	—
20 year (new)	Oct–93	Sep–96	10/1/1993	9/30/96
30 year	Feb–77	Sep–96	2/15/1977	9/30/96

constraining the historical scope of the analysis.

The second reason is that we need to take the overall slope of the yield curve into account. If the yield curve as a whole is nearly flat, the convexity bias will lead to an appreciable downward slope from the twenty-year to the thirty-year point; if it is steeply positive, and the same convexity bias is present, the curve may be flat, or even upward-sloping from twenty years to thirty years. Thus short bond yields are relevant to the analysis.

The third reason is that any analysis that relies too heavily on the yields of specific bonds is vulnerable to yield anomalies that affect those particular bonds. These may arise from temporary supply/demand or liquidity conditions, or may be more persistent (as in the case of the on-the-run ten-year note; see Carayannopoulos [1996]). Thus it is desirable to somehow take all bond yields into account in the analysis.

An important caveat is that very short-dated bond yields must be used with caution. These often show low correlations with other bond yields, and principal components analysis suggests that there are factors that specifically affect the short end of the yield curve, such as the formation of humps. This is consistent with market experience; the short end of the curve is often "distorted" when current monetary policy settings are perceived as aberrant.

Furthermore, market segmentation, the relatively low liquidity of short-maturity bonds, and interaction with the Eurodollar futures market all make the behavior of this part of the curve more complex. These phenomena, while interesting, are not relevant to the measurement of the convexity bias at the long end of the curve, and may well interfere with the analysis by introducing instability into the measurement of yield curve slope.

The approach adopted here is to use observed bond yields to construct a smooth "theoretical" yield curve, which eliminates yield anomalies and short-end aberrations as far as possible. Note that naive spline methods are ruled out, since the results would then be sensitive to the (arbitrary) choice of knot points; a further problem with splines is that they are inappropriate if too few bond yields are available. Instead, the smoothing process is based on an analysis of the theoretical structure of investor expectations about future money market yields.

In order to derive a formal expectations theory for the shape of the yield curve, we will: 1) define the "long rate" and the "short rate," and regard the forward rate curve as linking these two rates; 2) observe that the structure of investor expectations determines the functional form of the forward rate curve; 3) show how a bond market risk premium may be incorporated, and 4) show how a convexity adjustment may be incorporated. Although the technical details are very simple, it is worthwhile explaining the conceptual foundations in some detail, as this critically affects the interpretation of the results.

The long rate is defined to be the expected long-term return on bonds. That is, it is the "expected return on a bond with infinite duration," i.e., the limit of the expected return on a zero-coupon Treasury bond as the maturity approaches infinity. The long rate is not equal to the limiting zero-coupon yield, since yields incorporate a convexity adjustment. The long rate is determined by the market's expectations of long-run inflation and growth in the economy.

The short rate is defined to be the expected short-term return on bonds; this assumes that since the bond market is liquid and homogeneous, any difference in expected short-term returns between bonds will be arbitraged away. The short rate need not be equal to any specific bond yield, since the expected short-term return on a bond incorporates not just an income effect but also an expected price effect as reflected in observed forward bond prices. The short rate is determined by the market's expectations of imminent monetary policy, as influenced by near-term inflation and growth prospects; clearly, it is related to the short-term money market yield.

Note that the long rate l is not the consol yield: it is "longer" than a consol, because a consol has short-dated cash flows. Also note that the short rate s is not a money market yield, but must be interpreted as a bond market return that does not take account of a misalignment between bond market expectations and current monetary policy. For example, if the Fed funds rate is currently 5%, but the market expects that a 1% tightening is imminent, the short rate is likely to be closer to 6% than 5%.

Duffie and Kan [1993] remark that the short rate should be regarded as a limiting, rather than an actual, yield; a key insight due to Mason is that in constructing a model of bond yields, the short rate should be regarded as a limiting *bond* yield, factoring out the idiosyncrasies of the short-term money market.

Ignoring convexity adjustments for the moment, the short rate and the long rate can be thought of as being joined by an instantaneous forward rate curve g(t), so that $g(0) = s, \lim_{t \to \infty} g(t) = l$, and g(t) can be integrated appropriately to obtain any specific spot or forward bond yield. (Note that g(t) may be viewed as the forecast short rate at time t.) The hypothesis is that the shape of the forward rate curve is determined by the market's expectations of future bond market returns, adjusted for risk. A number of steps are required to spell this out.

First, assume that expected bond market returns are equal to expected money market yields; g(t) may then be interpreted as the expected money market yield at time t. Make the further assumption that the long-run money market yield $\lim_{t \to \infty} g(t) = l$ is equal to the sum of the long-run expected inflation rate and the long-run expected real interest rate, which must be true in equilibrium. Frankel [1995] shows that a wide variety of macroeconomic models — those including a money demand equation and a price adjustment equation — imply that the rationally expected path of money market yields g(t) must have the functional form:

$$g(t) = l + (s - l)e^{-\kappa t}$$

That is, assuming that investors form their expectations in accordance with a reasonable macroeconomic model — which may occur on a conscious or unconscious level — their expectations about future money market yields must take this form, at least beyond the near term. In the near term, as pointed out above, more detailed knowledge about likely monetary policy can result in a more complex structure of short-term rate expectations.

Note that the result holds even for macroeconomic models that allow future random disturbances to the level and trend of the money supply, provided these have expectation zero. Also note that we have *not* assumed that the economy is correctly described by a macroeconomic model of this kind; we have assumed only that investors (explicitly or implicitly) *believe* that it can be so described, at least when forming their expectations. Given the intuitive foundations of models of this kind, this appears to be an initially plausible assumption.

The coefficient κ is a rate of adjustment parameter, or a mean reversion parameter, which is derived from certain elasticity parameters in the macroeconomic model (which say, for example, how quickly prices respond to excess demand). Note that the value of κ is determined by the structure of the economy itself, and is unrelated to the fundamental determinants of interest rates such as the economic cycle. In forming expectations, then, investors would not expect κ itself to vary over time.

As explained above, s = g(0) should not be identified with the current money market yield; rather, it is derived from the bond market's prediction of the "imminent" money market yield, which factors out any perceived short-term aberrations in monetary policy. To put it differently: When the bond market perceives current money market yields as reflecting a dis-equilibrium state, s will reflect its judgment of where the short-term equilibrium money market yield should be, given its assessment of current economic conditions.

In practice, there is arguably a risk premium for holding Treasury bonds instead of money market securities (quite distinct from a convexity adjustment, which applies only to long bonds and has the opposite sign). For example, Ilmanen [1996] finds that although the historical risk premium is extremely difficult to estimate, it does appear to exist; moreover, it is approximately constant across bonds with maturities of three to twenty years.

Our framework can easily accommodate an expected bond market risk premium; one can add a risk premium to s and l. Note that the expected risk premium need not be constant by maturity, but it is constrained to vary smoothly with maturity, with the same rate of adjustment parameter κ.

The final step is to determine how to incorporate a convexity adjustment to obtain an assumed forward rate curve f(t). To do this, we derive a rule of thumb that is commonly employed in the swap market. Consider an arbitrary bond without embedded options. If the yield of the bond at some fixed future date t is y, write the bond price as B(y). For example, the current t-forward price of the bond is:

$$B_{fwd} = B(y_{fwd})$$

where y_{fwd} is the current t-forward yield.

Now, the expected bond price at time t is just $E[B] = B_{fwd}$. If the bond yield is volatile, however, it is not the case that $E[y] = y_{fwd}$; in fact, the difference between the forward yield and the expected future yield is precisely the convexity adjustment.

If the actual yield of the bond at time t is y rather than y_{fwd}, then we can write:

$$B(y) \approx B(y_{fwd}) + D(y_{fwd} - y) + \frac{1}{2}C(y_{fwd} - y)^2$$

where D is the duration of the forward bond position, and C is its convexity. Taking expected values, this gives us:

$$B_{fwd} \approx B_{fwd} + D(y_{fwd} - E[y]) + \frac{1}{2}CE[(y_{fwd} - y)^2]$$

Now $E[(y_{fwd} - y)^2] = \sigma^2 t$, where σ is the expected basis point volatility of long bond yields. Therefore, by cancellation:

$$E[y] - y_{fwd} \approx \frac{1}{2}(C/D)\sigma^2 t$$

In particular, the convexity adjustment for a forward rate adjustment is approximately $^1/_2\sigma^2 t^2$. This gives the relationship between the expected future short rate and the forward rate at a future time t.

To summarize, then, our expectations theory states that the forward rate curve has the functional form:

$$f(t) = l + (s - l)e^{-\kappa t} - \frac{1}{2}\sigma^2 t^2$$

where σ is the expected absolute volatility of long bond yields. The four parameters need to be estimated from a given set of Treasury bond yields, i.e., a given yield curve. Note that the presence of the mean reversion parameter means that this model does not fit into the framework of Duffie and Kan [1993].

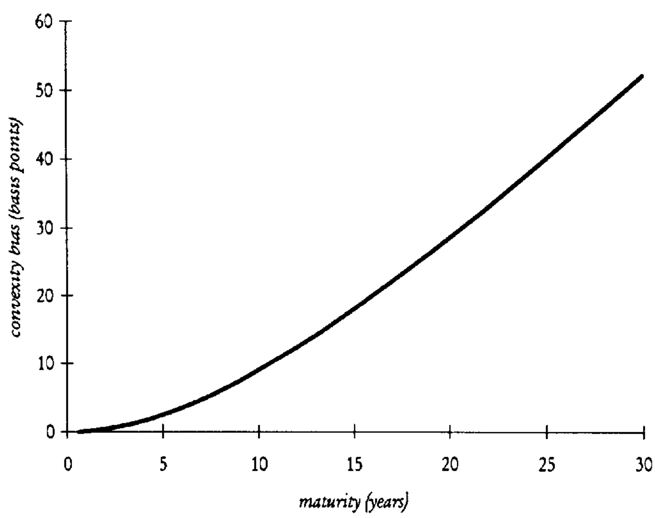
This formulation has the undesirable feature that very long forward rates become negative, causing very long forward bond prices to explode, but this is not a significant problem in practice. Taking l = 6% (i.e., a historical low for the post-1970 period) and σ = 0.75% (corresponding to 12.5% proportional volatility, a high estimate for long-term volatility), forward rates only become negative at a forward date of forty-six years, and the thirty-year forward rate is approximately 3.5%. Thus, it may be hoped that the model gives reasonable answers when applied to Treasury bonds, although more sophisticated methods would have to be devised to analyze hundred-year bonds.

More reasonable asymptotic behavior could be obtained by assuming that long bond yields are mean-reverting. Unfortunately, this introduces another parameter that is extremely difficult to estimate — much more so than κ, which is itself hard to estimate. Moreover, the mean reversion time scale would necessarily be long (for example, long bond yields broadly trended upward from 1953 to 1982). Thus, the practical impact on the analysis of Treasury bond yields would be minimal; the results in this article would not be affected.

The point is that even though our formulation has implausible asymptotic properties, it is realistic for bonds with maturities of up to thirty years; thus it may still do a good job of modeling the investor expectations that determine Treasury bond yields. In fact, as a model of the structure of investor expectations, it is probably more plausible than one based on a mathematically sophisticated term structure theory with a larger number of unobservable parameters, even if the latter is mathematically more consistent. The test, of course, is to fit the model to historical Treasury yields and see whether the results are meaningful.

Exhibit 2 graphs the convexity adjustment versus bond maturity, based on a rule of thumb and assuming an absolute yield volatility of 80 bp per year. Note that the model can be fitted to a single yield curve, i.e., to cross-sectional data. We do not need to make any assumptions about the time series properties of Treasury yields in order to estimate implicit volatility

EXHIBIT 2 ■ Theoretical Convexity Bias Versus Maturity, 80 bp Annual Volatility



(or the mean reversion coefficient) using the model. Besides making the methodology much more robust, this also means that much more information can be extracted from the data, in the form of a historical time series of parameter estimates.

II. ESTIMATION PROCEDURE

The forward rate curve can be integrated to obtain formulas for the discount function and par yield curve:

$$PV(t) = \exp\left(-lt - (s - l)\kappa^{-1}(1 - e^{-\kappa t}) + \frac{1}{6}\sigma^2 t^3\right)$$

$$\text{Theoretical n-Year CMT Yield} = \frac{2[1 - PV(n)]}{\sum_{i=1}^{2n} PV(i/2)}$$

In principle, then, the estimation problem is simple. Given a set of constant-maturity Treasury (CMT) yields, find the l, s, σ, and κ that give the best fit to those yields. The first question is: What is the meaning of "best fit"? I adopt the normal procedure of initially using χ^2 estimation, and then reviewing the distribution of errors to determine whether this is in fact a reasonable thing to do.

It is helpful to make some observations before proceeding. First, provided the model of investor expectations is realistic, one would not expect serious outliers. Deviations from the model may be thought of as indicating anomalous liquidity or supply/demand conditions, and under this interpretation a discrepancy no larger than 5–15 basis points (and generally less than this for longer bonds) would be expected. Thus it should not be necessary to use more robust statistical estimators.

As noted above, however, the model need not be realistic for short-dated bonds. Thus, if one-year CMT yields are included in the data set, much larger discrepancies would be expected under certain circumstances — for example, at times when there is a sharp divergence between Fed policy and the market's views. It should also be noted that, in economic terms, the same basis point error is less significant for a shorter than for a longer bond. It might thus be possible that somewhat larger discrepancies could arise for shorter bonds.

The solution adopted is to weight basis point errors by bond duration, i.e., to minimize price errors rather than yield errors. This turns out to give empirically good results for two– to thirty-year bond yields. It gives the one-year CMT minimal weight without discarding it from the data set altogether.

The study implements χ^2 estimation using the routines provided in the Optimization Toolbox of MATLAB®; see Grace [1994] (the initial reference for the algorithms referred to below). These prove to be satisfactorily robust and efficient, and have the advantage of providing useful diagnostics; for example, they report condition numbers of any severely ill-conditioned matrices that arise.

It might seem most natural to regard this as a constrained non-linear optimization problem, where natural constraints are l, s, σ, and $\kappa > 0$. Negative long-term rates are economically unreasonable; negative short-term rates have existed only rarely and for brief periods, under very special circumstances; and negative volatilities or mean reversion coefficients make no sense. Efficient methods for solving constrained problems have been developed relatively recently.

Unfortunately, the presence of κ as a fitted parameter in the model specification considerably complicates the estimation process. Bond yields are affine functions of l, s, and σ^2 but not κ. For very small and very large κ, derivatives involving κ are badly behaved, leading to severe poor conditioning in methods that make use of numerically computed gradients. The κ term also makes the behavior of any such algorithm overly sensitive to the choice of initial guess. For this reason, it is decided to solve the unconstrained problem first, using the robust Nelder-Mead simplex algorithm; see Press et al. [1992].

As expected, this algorithm converges in all cases but often gives meaningless results, including negative volatilities and/or absurdly large values presumably resulting from an unlucky choice of initial guess. When convergence to a meaningful set of values is achieved, κ generally lies between 0.1 and 2, and σ is generally less than 2%. Thus, this exercise establishes that it is valid to impose the constraints l, s > 0, $0.1 \le \kappa \le 2$, and $0\% \le \sigma \le 2\%$. This has the effect of restricting the search to economically reasonable values, and of excluding regions where the κ derivatives are badly behaved.

The next step is to solve the new constrained problem using a sequential quadratic programming algorithm, which employs an active set strategy at each iteration. Because of the constraints on κ, this algorithm converges rapidly in all cases, and yields intuitively reasonable long and short rates (as illustrated in Exhibit 3). Furthermore the time series of fitted long and short rates do not exhibit more volatility than the time series of observed long and short bond yields, implying that the estimation process is stable.

Exhibit 4 shows histograms of basis point errors for the five-year and ten-year bond. Note that except in the case of the one-year CMT yield, errors are generally small (less than 10 bp 90% of the time); also, errors seem to be approximately normally distributed, which suggests that it is reasonable to employ χ^2 estimation. There is no bias except for the ten-year CMT yield, which is generally overestimated by the model; this, however, is consistent with the findings of Carayannopoulos [1996] regarding the systematic overpricing of the ten-year Treasury note.

Exhibit 5 shows the time series behavior of the error for the ten-year and thirty-year CMT yield. The error appears to be a stationary process that is uncorrelated with the level or slope of the yield curve, or with any other obvious economic variable. It may be concluded that the yield curve model we have adopted is a realistic one.

Interestingly, the distribution of estimates of κ is approximately bimodal. Values cluster roughly around 0.1 (the lower constraint) and 0.5 (using an initial guess of 0.4), with the latter peak significantly more impor-

EXHIBIT 3 ■ Model Versus Actual Yields: Basis Point Errors

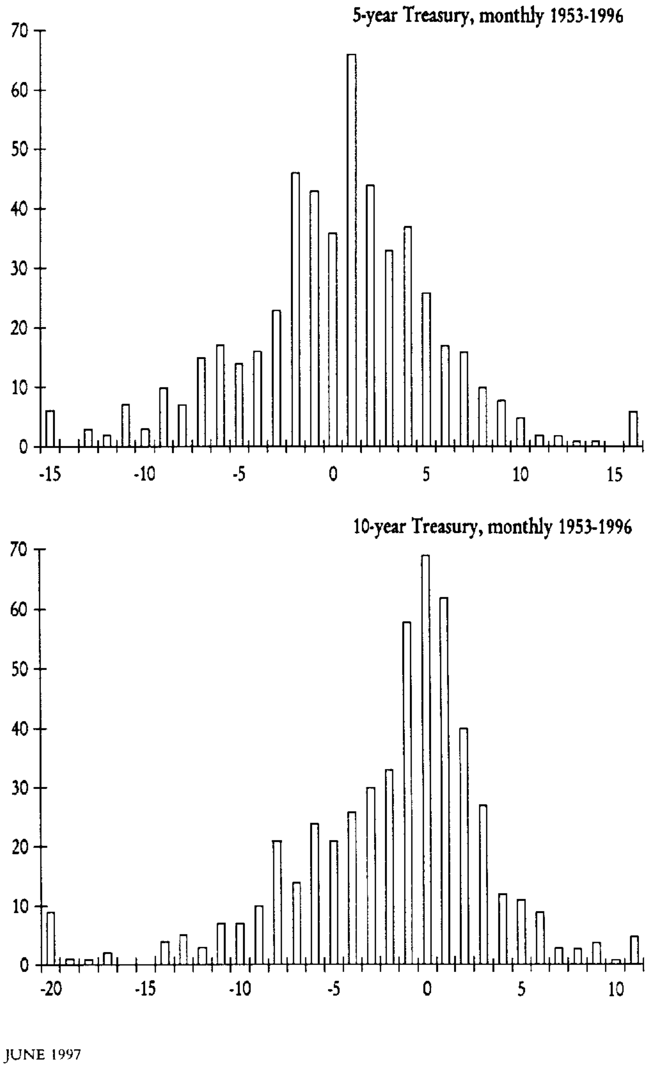


EXHIBIT 4 ■ Model Basis Point Errors Over Time

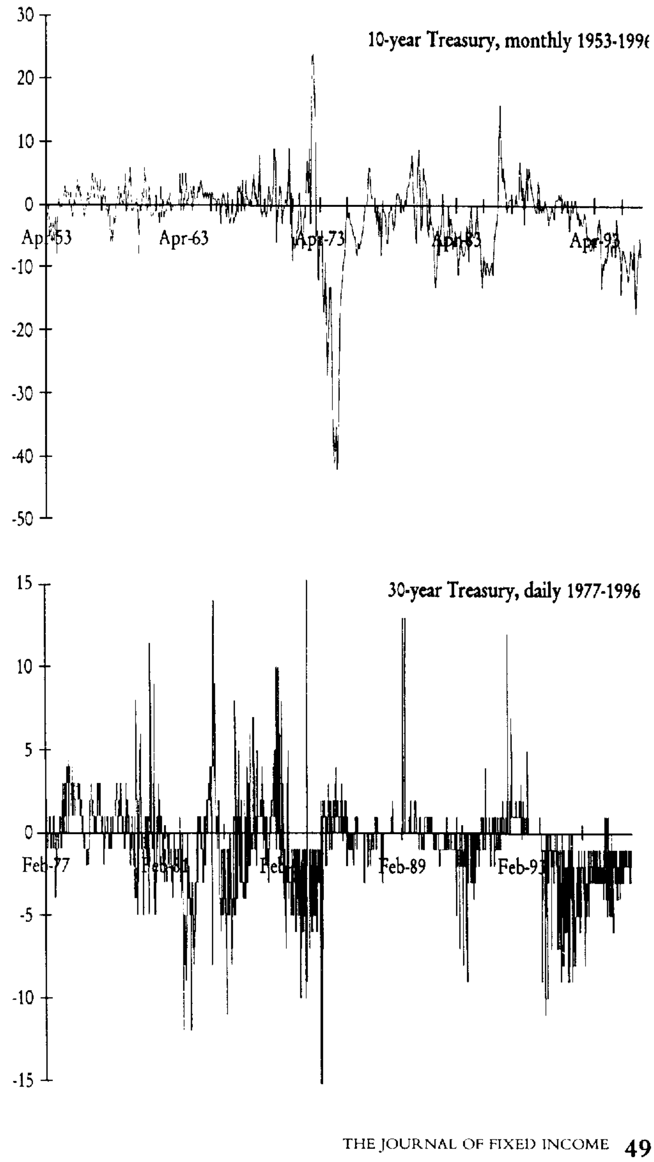
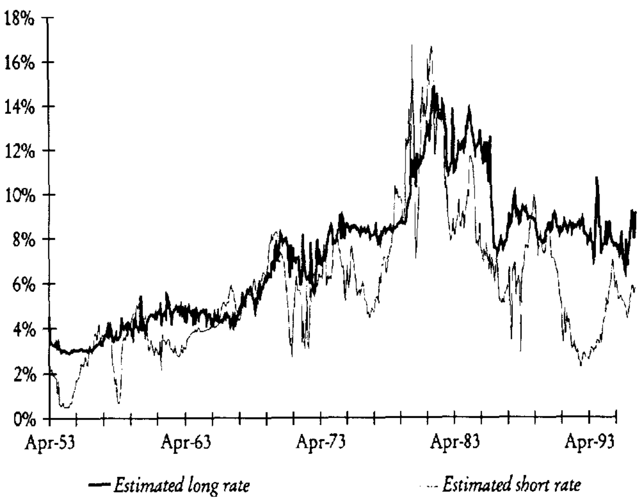


EXHIBIT 5 ■ Long and Short Rates, Monthly Estimates 1953–1996



tant. Examination of the time series shows that, apart from short-term spikes, κ moves significantly away from 0.5 only in the periods 1959–1969, when it was generally close to 0.1, and in the period 1991–1996, when it fluctuated between 0.1 and 0.75; see Exhibits 6 and 7.

It is possible to view κ as a curvature parameter; the predicted degree of curvature of the yield curve will be roughly proportional to $\kappa(1 - s)$. This means that estimates of κ are primarily influenced by three-year to seven-year bond yields. By contrast, estimates of σ are mainly influenced by ten-year to thirty-year bond yields. Thus, if we are interested in estimating σ, then variations in κ should not make too much difference — although in principle all four quantities are coupled.

In the light of this observation, the constrained problem with $\kappa = 0.5$ is solved using a standard Levenberg-Marquardt algorithm. This converges rapidly in all cases, and always gives meaningful estimates for the long and short rates. In most cases it also gives estimates for implicit volatility consistent with the previous estimates, although there is some divergence when the yield curve is very steep. In the period 1993–1996, however, the estimates for σ (setting $\kappa = 0.5$) are about 50% of the estimates for σ (allowing κ to be a fitted parameter). That is, when the "actual" value of κ diverges significantly from 0.5, as it did during most of this period, a procedure that assumes that $\kappa = 0.5$ can generate consistently inaccurate estimates for σ.

Thus, as one might expect, the rate of mean reversion should not generally have a significant impact on the pricing of long bonds, provided that a ballpark value is assumed (but 0.5 is not always a suitable ballpark value). That is, the problem of estimating implicit volatility is more robust than it might have appeared at first. Estimating the mean reversion coefficient as well as the other parameters using cross-sectional data is still

EXHIBIT 6 ■ Mean Reversion Coefficient, Monthly Estimates 1953–1996

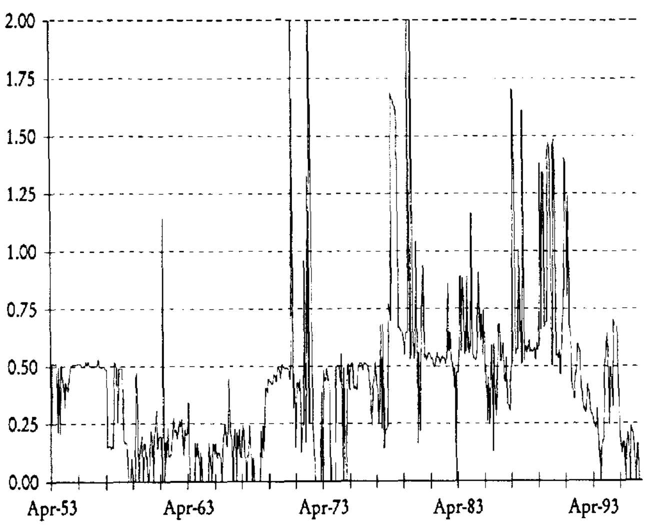


EXHIBIT 7 ■ Implicit Volatility, Monthly Estimates 1953–1996

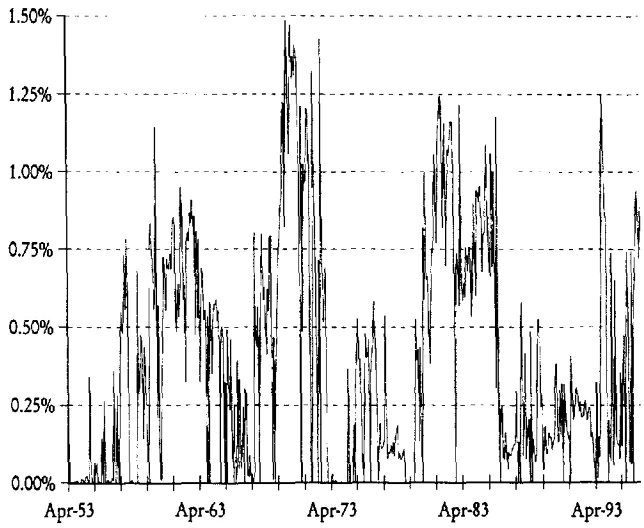
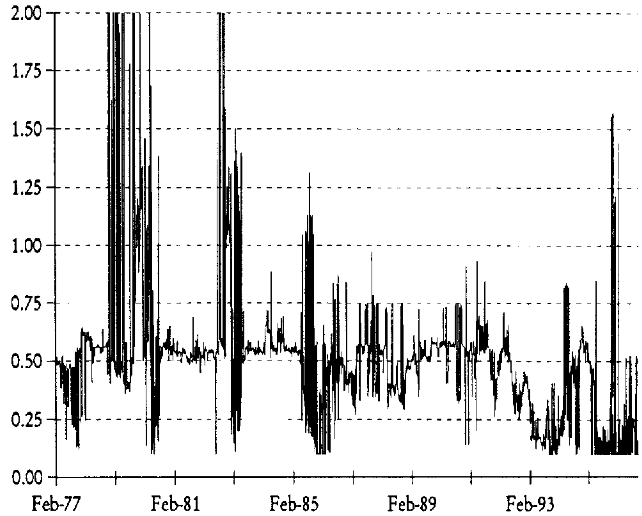


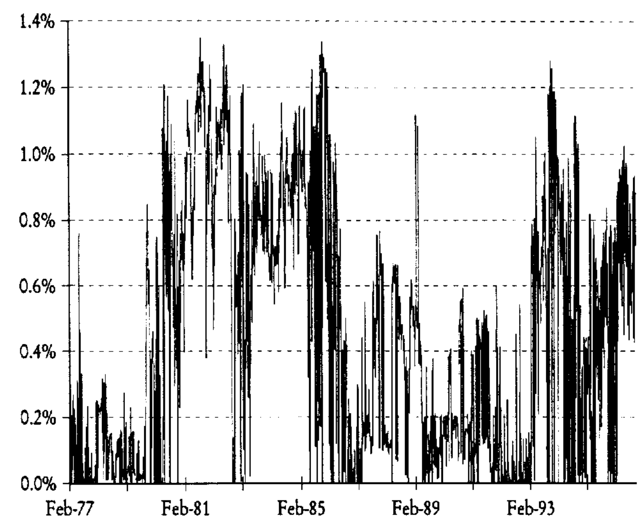
EXHIBIT 8 ■ Mean Reversion Coefficient, Daily Estimates 1977–1996



likely to give more reliable results.

It is perhaps fortunate that accurate estimates for κ are not required, since estimating this quantity is an inherently unstable problem — partly because, when the yield curve is flat or nearly flat, almost any value of κ will give rise to the same theoretical yields. Note that the estimation procedure adopted here, which weights errors by duration and hence gives a high weight to long bond yields, probably does not generate the most accurate possible estimates of κ.

EXHIBIT 9 ■ Implicit Volatility, Daily Estimates 1977–1996



III. RESULTS, 1953–1996 U.S. TREASURY DATA

Exhibit 8 shows the monthly estimates of implicit volatility for the period 1953–1996. Exhibit 9 shows the daily estimates for the period February 1977 – September 1996. All results are expressed in percent per year. As one would expect, the results are quite consistent.

EXHIBIT 10 ■ Implicit Volatility Regimes

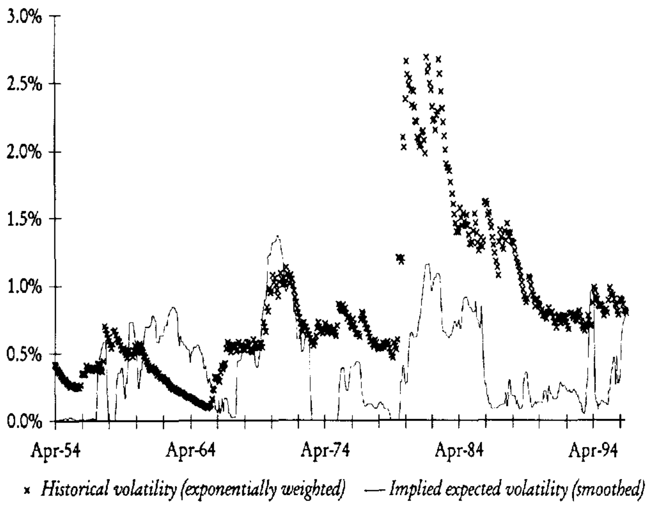
Period	Regime	Approximate Annual Volatility
1954–1957	Off	–
1958–1965	On	70 bp
1966–1967	Off	–
1968–1972	On	100 bp
1973–1979	(Mostly) off	–
1980–1983	On	80 bp
1984–1993	Off	–
1994–1996	(Mostly) on	80 bp

The major observation is that implicit volatility is not constant, nor does it seem to be a continuously varying random quantity. Instead, it seems to be either switched "on" or "off"; there appear to be two distinct regimes (see Exhibit 10).

Exhibit 11 compares implicit volatility with historical ten-year bond yield volatility, both expressed in percent per year. (Implicit volatility is smoothed by taking an annual median, to make the graph easier to read.) Historical volatility is exponentially weighted with a decay factor of 0.9, and is thus a reasonably long-term measure, while still responding quickly to volatility shocks.

Implicit volatility can be switched "off" even while bond yields are still quite volatile, but a volatility *shock* tends to switch implicit volatility "on." The most recent examples of this occurred in 1979–1980 and 1993–1994.

EXHIBIT 11 ■ Implicit Volatility Versus Historical Volatility



The other interesting observation is that, in basis point terms, historical volatility tends to be higher when yields are higher, while the level of implicit volatility (when it is "on") seems to be independent of absolute yield levels. Remembering that implicit volatility is meant to encapsulate the market's volatility expectations over long time frames, during which outright yields will vary greatly, it seems a priori reasonable that the fact that yields are currently high or low should not unduly bias implicit volatility, and this does in fact appear to be the case.

If one believes that the volatility of yields is proportional to the outright level of yields, an implicit volatility of 70–80 bp is consistent with the hypothesis that the market expects long-term proportional volatility to be around 11%, and that long bond yields will fluctuate around 7.5%, i.e., are mean-reverting. But the evidence for this interpretation is not very strong.

Although complete data are not available for this study, it appears that implicit volatility is not closely related to long-dated OTC implied volatilities. This is not surprising, for the reasons mentioned in the introduction.

Finally, the mean reversion coefficient itself seems to exhibit a kind of regime-switching behavior. During the periods 1953–1957, 1972–1992, and 1994, it appears to fluctuate around 0.5, while during 1958–1971 and 1993–1996 (excepting 1994), it appears to fluctuate around 0.1–0.2. A very tentative interpretation is that a low value of κ appears to coincide with long-term expectations of strong growth coupled with low inflation and a belief that the business cycle no longer applies — these are clearly secular, rather than cyclical, shifts in attitudes.

It would be unwise to attempt to draw any definitive conclusions, because the estimation procedure is not optimized to generate the most accurate possible estimates of κ. A more refined procedure would probably employ different weights, and should make use of the yields of traded bonds, rather than interpolated constant-maturity Treasury yields.

IV. CONCLUSIONS

A series of studies beginning with Litterman and Scheinkman [1991] have attempted to isolate the major factors that drive shifts in the yield curve. The tool of choice has been principal components analysis. The two dominant factors consistently turn out to be a parallel shift in the yield curve and a shift in yield curve slope, with a curvature shift generally emerging as the third factor. These may be interpreted, in terms of our model, as changes in l, (l – s), and κ.

This has some implications for risk management. In the discussion following Brown and Schaefer [1995], it is noted that two or three factors appear to suffice for attempting to capture yield curve risk. My work suggests the existence of an additional source of yield curve risk — fluctuations in implicit volatility, σ — which should be taken into account in the risk monitoring process.

This source of risk may not be apparent from a principal components analysis, because, as we have seen, σ does not exhibit small random fluctuations like a "typical" random quantity, but instead seems to be switched "on" or "off." Furthermore, studies tend to use data sets whose observations are much more closely clustered at short maturities, thereby reducing the statistical weight given to observed fluctuations in long bond yields. Thus a short-end "hump" factor is much more likely to be detected, and will appear relatively more important than a "convexity bias" factor driven by changes in implicit volatility.

Volatility risk is, of course, very familiar from the interest rate derivatives markets. In fact, in those markets, where volatility shows a distinct term structure, it is necessary to subdivide volatility risk by (for example) separating exposure to short-dated and long-dated volatility.

In this study, by contrast, it does not make practical sense to introduce more than one volatility parameter. Although theoretically attractive, this would make the estimation process unstable, and would probably lead to meaningless results. (Note that mean reversion implies that physical bond yield volatilities have some term structure anyway; i.e., the model is already realistic in that it does not predict flat volatilities.)

It would be interesting to determine how implicit volatility risk could be hedged. This question might be studied in the framework of a model of the stochastic evolution of l, (l – s), κ, and σ. For example, the regime-switching behavior of κ and σ could be modeled in a continuous framework using noisy Van der Pol oscillators. As various technical obstacles arise, this line of investigation must be postponed.

On a more practical note, the present line of investigation has potential applications to fixed-income portfolio management. For example, it could be argued that when implicit volatility is low, as it was in 1987–1993 except for brief periods, thirty-year bonds are undervalued relative to shorter bonds. This has implications for an active portfolio strategy with a relatively long time horizon; in effect, in periods such as these, long-dated long bond volatility can be bought very cheaply.

There are a number of reasons why this strategy is not entirely obvious, which may explain why such "anomalies" can persist for long periods of time. In particular, note that one cannot exploit the opportunity simply by executing a duration-matched switch from, say, twenty-year bonds into thirty-year bonds, as this would change the exposure of the portfolio to a shift in yield curve slope. Instead, one must adopt a neutral strategy that leaves slope risk as well as duration unaffected (see, e.g., Willner [1996]); a twenty-year position must be rebalanced into ten-year and thirty-year holdings, with a correspondingly smaller convexity pickup. Thus the trade is worth executing only in volume. Incidentally, this strategy ignores curvature risk; κ-immunization is possible via "condor" rather than "butterfly" trades, but is unwieldy and subject to additional execution risk.

Because these "anomalies" may persist for a long time, and because funding short bond positions is costly, it probably does not make sense to execute this kind of convexity arbitrage on a leveraged basis. Thus one would not necessarily expect hedge fund activity or proprietary trading to drive such anomalies away.

Alternatively, rather than regarding low implicit volatility as an anomaly, one might choose to interpret it as suggesting a risk premium for thirty-year bonds over and above the overall bond market risk premium (versus money market returns). During 1987–1993, this implied risk premium would have averaged 40–50 bp. But it is difficult to see why this risk premium should vanish at points when long bond yields become very volatile, which is what it seems to do.

One would also be forced to conclude that long bonds had a negative risk premium in the early 1960s. Although this is possible — in recent experience they have certainly been much less volatile than short rates — it is somewhat counterintuitive.

A final practical observation is that the shape of the long end of the curve — which is determined mainly by σ — is, to a large degree, independent of the shape of the middle part of the curve — which is determined mainly by κ, although the formulation of Willner [1996] may perhaps be more useful for practical applications. That is, investment managers may be justified in assuming that long bond "value" and mid-range bond "value" are not too closely coupled, and in adopting different methods to analyze different parts of the curve.

This work potentially sheds some light on alternative theories of the term structure, in particular, the preferred habitat theory. It may be possible to relate regime switching to structural changes in the fixed-income market triggered by volatility. This requires a detailed institutional analysis, and is beyond the scope of this article.

We conclude by noting the usefulness of examining a long historical time period when carrying out any analysis of bond market behavior. In the present case, the data exhibit noticeable patterns over time frames longer than an economic cycle, presumably related to secular changes in the structure of the economy or the U.S. bond market. Furthermore, an analysis of (say) 1990–1996 data alone would be potentially misleading. This observation should be borne in mind when extending the analysis to bond markets other than the United States.

ENDNOTE

The author thanks Charles Belvin, Richard Mason, and Michael Shearer for fruitful discussions. An important

initial motivation was an attempt to determine what is missing from a formal expectations theory advanced by Richard Mason.

REFERENCES

Brown, R., and S. Schaefer. "Interest Rate Volatility and the Shape of the Term Structure." In S. Howison, F. Kelly, and P. Wilmott, eds., *Mathematical Models in Finance*. London: Chapman and Hall, 1995.

Burghardt, G., and B. Hoskins. "A Question of Bias." *Risk,* March 1995.

Carayannopoulos, P. "A Seasoning Process in the U.S. Treasury Bond Market: The Curious Case of Newly Issued Ten-Year Notes." *Financial Analysts Journal,* January/February 1996.

Duffie, D., and R. Kan. "A Yield-Factor Model of Interest Rates." Working paper, 1993.

Frankel, J. *Financial Markets and Monetary Policy*. Cambridge: MIT Press, 1995.

Grace, A. *Optimization Toolbox For Use with MATLAB®*. The MathWorks Inc., 1994.

Ilmanen, A. "Does Duration Extension Enhance Long-Term Expected Returns?" *Journal of Fixed Income,* September 1996.

——."Market Rate Expectations and Forward Rates," *Journal of Fixed Income,* September 1996.

Litterman, R., and J. Scheinkman. "Common Factors Affecting Bond Returns." *Journal of Fixed Income,* June 1991.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C: The Art of Scientific Computing* (2nd ed.) Cambridge: Cambridge University Press, 1992.

Willner, R. "A New Tool for Portfolio Managers: Level, Slope, and Curvature Durations." *Journal of Fixed Income,* June 1996.