The Journal of
FIXED INCOME
VOLUME 7 NUMBER 2 www.iijfi.com SEPTEMBER 1997

DYNAMICS OF THE SHAPE OF THE YIELD CURVE

ANTTI ILMANEN AND RAY IWANOWSKI

DYNAMICS OF THE SHAPE OF THE YIELD CURVE

ANTTI ILMANEN AND RAY IWANOWSKI

How can we interpret the shape (steepness and curvature) of the yield curve *on a given day*? And how does the yield curve evolve *over time*? In this article, we examine these two broad questions about the yield curve behavior. Yield curve shape reflects the market's rate expectations, required bond risk premiums, and convexity bias. We discuss various economic hypotheses and empirical evidence about the relative roles of these three determinants in influencing the curve steepness and curvature. We also discuss term structure models that describe the evolution of the yield curve over time and summarize relevant empirical evidence.

The key determinants of the curve steepness are the market's rate expectations and the required bond risk premiums. The pure expectations hypothesis assumes that all changes in steepness reflect the market's shifting rate expectations, while the risk premium hypothesis assumes that changes in steepness reflect only changing bond risk premiums. In reality, rate expectations *and* required risk premiums do influence the curve slope.

Historical evidence suggests that above-average bond returns, and not rising long rates, are likely to follow abnormally steep yield curves. Such evidence is inconsistent with the pure expectations hypothesis and may reflect time-varying bond risk premiums. Alternatively, the evidence may represent irrational investor behavior and the long rates' sluggish reaction to news about inflation or monetary policy.

The determinants of the yield curve's curvature have received less attention. It appears that curvature varies primarily with the market's curve reshaping expectations. Flattening expectations make the yield curve more concave (humped), and steepening expectations make it less concave or even convex (inversely humped).

ANTTI ILMANEN is the European bond strategist at Salomon Brothers in London, England.

RAY IWANOWSKI is a vice president at Goldman Sachs Asset Management in New York.

It seems unlikely, however, that the average concave shape of the yield curve results from systematic flattening expectations. More likely, the shape reflects the convexity bias and the required return differential between barbells and bullets. If convexity bias were the only reason for the concave average yield curve shape, one would expect a barbell's convexity advantage to exactly offset a bullet's yield advantage, in which case the duration-matched barbell and bullet would have the same expected returns. Historical evidence suggests otherwise: In the long run, bullets have earned slightly higher returns than duration-matched barbells. That is, the risk premium curve appears to be concave rather than linear in duration.

We discuss plausible explanations for the fact that investors, in the aggregate, accept lower expected returns for barbells than for bullets; the barbell's lower return volatility (for the same duration); the tendency of a flattening position to outperform in a bearish environment; and the insurance characteristic of a positively convex position.

In our work on yield curve evolution, we describe some empirical characteristics of the yield curve behavior that are relevant for evaluating various term structure models. The models differ in their assumptions regarding the expected path of short rates (degree of mean reversion); the role of a risk premium; the behavior of the unexpected rate component (whether yield volatility varies over time, across maturities or with the rate level); and the number and identity of factors influencing interest rates.

For example, the simple model of parallel yield curve shifts is consistent with no mean reversion in interest rates and constant risk premiums over time. Across bonds, the assumption of parallel shifts implies that the term structure of basis point yield volatilities is flat and that rate changes are perfectly correlated (and, thus, driven by one factor).

Empirical evidence suggests that short rates exhibit quite slow mean reversion; required risk premiums vary over time; yield volatility varies over time (partly related to the yield level); the term structure of basis point yield volatilities is typically inverted or humped; and rate changes are highly but not perfectly correlated across the curve.

I. HOW SHOULD WE INTERPRET THE YIELD CURVE STEEPNESS?

The steepness of the yield curve primarily reflects the market's rate expectations and required bond risk premiums because the third determinant, convexity bias, matters mainly at the long end of the curve. A particularly steep yield curve may be a sign of prevalent expectations for rising rates, abnormally high bond risk premiums, or some combination of the two. Conversely, an inverted yield curve may be a sign of expectation of declining rates, negative bond risk premiums, or a combination of declining rate expectations and low bond risk premiums.

We can map statements about the curve shape to statements about the forward rates (ignoring convexity effects). When the yield curve is upward sloping, forwards "imply" rising rates. The implied forward (par and spot) yield curves show the break-even levels of future yields that would cause capital losses that exactly offset the longer bonds' yield advantage over the riskless short bond. That is, if these implied forward yield curves are subsequently realized, all bonds will earn the same holding-period return.

Because expectations are not observable, we do not know with certainty the relative roles of rate expectations and risk premiums. It may be useful to examine two extreme hypotheses that claim that the forwards reflect only the market's rate expectations or only the required risk premiums. If the pure expectations hypothesis holds, the forwards reflect the market's rate expectations, and the implied yield curve changes are likely to be realized (that is, rising rates tend to follow upward-sloping curves, and declining rates tend to succeed inverted curves). If the risk premium hypothesis holds, the implied yield curve changes are not likely to be realized (that is, high excess bond returns tend to follow upward-sloping curves, and low excess bond returns tend to succeed inverted curves).

Empirical Evidence

To evaluate these hypotheses, we compare implied forward yield changes (which are proportional to the steepness of the forward rate curve) to subsequent average realizations of yield changes and excess bond returns.[1] In Exhibit 1, we report 1) the average spot yield curve shape; 2) the average of the yield changes that the forwards imply for various constant-maturity spot rates over a three-month horizon; 3) the average of realized yield changes over the subsequent three-month horizon; 4) the difference between 2) and 3), or the average "forecast error" of the forwards; and 5) the estimated correlation coefficient between the implied yield

EXHIBIT 1 ■ Evaluating the Implied Treasury Forward Yield Curve's Ability to Predict Actual Rate Changes, 1968-1995

	3 Mo.	6 Mo.	9 Mo.	1 Yr.	2 Yr.	3 Yr.	4 Yr.	5 Yr.	6 Yr.
Mean Spot Rate	7.04	7.37	7.47	7.57	7.86	8.00	8.12	8.25	8.32
Mean Implied Rate Change	0.65	0.32	0.27	0.23	0.14	0.12	0.11	0.08	0.07
Mean Realized Rate Change	0.003	0.001	0.000	0.000	0.001	0.001	0.001	0.002	0.002
Mean Forecast Error	0.65	0.32	0.27	0.23	0.14	0.12	0.11	0.08	0.07
Correlation Between Implied and Realized Rate Changes	–0.04	–0.08	–0.10	–0.08	–0.10	–0.13	–0.13	–0.12	–0.13

Notes: The spot yield curves are estimated on the basis of Treasury on-the-run bill and bond data using a relatively simple interpolation technique. Note that the longest spot yield we examine corresponds roughly to the yield of a ten-year OTR coupon bond whose average duration is slightly longer than six years during our sample period. (Given that the use of such synthetic bond yields may add some noise to the analysis, we have ensured that our main results also hold for yield curves and returns of actually traded bonds, such as OTR coupon bonds and maturity subsector portfolios.) The implied rate change is the difference between the constant-maturity spot rate that the forwards imply in a three-month period and the current spot rate. The implied and realized spot rate changes are computed over a three-month horizon using (overlapping) monthly data. The forecast error is their difference.

changes and the realized yield changes over three-month horizons. We use overlapping monthly data between January 1968 and December 1995 — deliberately selecting a long neutral period in which the beginning and ending yield curves are very similar.

Exhibit 1 shows that, on average, the forwards imply rising rates, especially at short maturities — simply because the yield curve tends to be upward sloping. The rate changes that would offset the yield advantage of longer bonds, however, have not materialized, on average, leading to positive forecast errors. Our unpublished analysis shows that this conclusion holds over longer horizons than three months and over various subsamples, including flat and steep yield curve environments. The fact that the forwards tend to imply too high rate increases is probably caused by positive bond risk premiums.

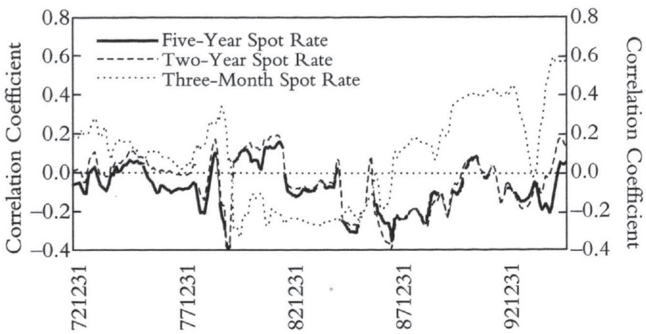
The last row in Exhibit 1 shows that the estimated correlations of the implied forward yield changes with subsequent yield changes are negative. These estimates suggest that, if anything, yields tend to move in the direction opposite from the direction that the forwards imply. Intuitively, small declines in long rates have followed upward-sloping curves, on average, thus augmenting the yield advantage of longer bonds (rather than offsetting it). Conversely, small yield increases have succeeded inverted curves, on average. The big bull markets of the 1980s and 1990s occurred when the yield curve was upward sloping, while the big bear markets in the 1970s occurred when the curve was inverted. We stress, however, that the negative correlations in Exhibit 1 are quite weak; they are not statistically significant.[2]

Many market participants believe that the bond risk premiums are constant over time and that changes in the curve steepness, therefore, reflect shifts in the market's rate expectations. The empirical evidence in Exhibit 1 and in many earlier studies contradicts this conventional wisdom. Historically, steep yield curves have been associated more with high subsequent excess bond returns than with ensuing bond yield increases.[3]

One may argue that the historical evidence in Exhibit 1 is no longer relevant. Perhaps investors forecast yield movements better nowadays, partly because they can express their views more efficiently with easily tradable tools, such as Eurodeposit futures. Some anecdotal evidence supports this view. Unlike the earlier yield curve inversions, the most recent inversions (1989 and 1995) were quickly followed by declining rates.

If market participants actually are becoming better forecasters, subperiod analysis should indicate that the implied forward rate changes have become better predictors of the subsequent rate changes. That is, the rolling correlations between implied and realized rate changes should be higher in recent samples than earlier. In Exhibit 2, we plot such rolling correlations, demonstrating that the estimated correlations have

EXHIBIT 2 ■ 60-Month Rolling Correlations Between the Implied Forward Rate Changes and Subsequent Spot Rate Changes, 1968-1995



Notes: The Treasury spot yield curves are estimated on the basis of OTR bill and bond data. The implied rate change is the difference between the constant-maturity spot rate that the forwards imply in a three-month period and the current spot rate. The implied and realized spot rate changes are computed over a three-month horizon using (overlapping) monthly data.

increased somewhat over the past decade.

In Exhibit 3, we compare the forecasting ability of Eurodollar futures and Treasury bills/notes in the 1987-1995 period. The average forecast errors are smaller in the Eurodeposit futures market than in the Treasury market, reflecting the flatter shape of the Eurodeposit spot curve (and perhaps the systematic "richness" of the shortest Treasury bills). The correlations between implied and realized rate changes suggest by contrast that the Treasury forwards predict future rate changes slightly better than the Eurodeposit futures do.

A comparison with the correlations in Exhibit 1 (the long sample period) shows that the front-end Treasury forwards, in particular, have become much better predictors over time. For the three-month rates, this correlation rises from −0.04 to 0.45, while for the three-year rates, the correlation rises from −0.13 to 0.01. Thus, recent evidence is more consistent with the pure expectations hypothesis than the data in Exhibit 1, but these relations are so weak that it is too early to tell whether the underlying relation actually has changed. Anyway, even the recent correlations suggest that bonds longer than a year tend to earn their rolling yields.

EXHIBIT 3 ■ Evaluating the Implied Eurodeposit and Treasury Forward Yield Curve's Ability to Predict Actual Rate Changes, 1987-1995

Eurodeposits	**3 Mo.**	**6 Mo.**	**9 Mo.**	**1 Yr.**	**2 Yr.**	**3 Yr.**	**4 Yr.**	**5 Yr.**	**6 Yr.**
Mean Spot Rate	6.32	6.40	6.48	6.58	6.98	—	—	—	—
Mean Implied Rate Change	0.16	0.18	0.19	0.20	0.20				
Mean Realized Rate Change	−0.02	−0.02	−0.02	−0.02	−0.03				
Mean Forecast Error	0.18	0.20	0.21	0.22	0.23				
Correlation Between Implied and Realized Rate Changes	0.39	0.18	0.11	0.06	0.02				
Treasuries									
Mean Spot Rate	5.67	5.90	6.01	6.13	6.64	6.86	7.07	7.29	7.41
Mean Implied Rate Change	0.47	0.30	0.27	0.28	0.19	0.16	0.15	0.12	0.11
Mean Realized Rate Change	−0.01	−0.01	−0.01	−0.01	−0.02	−0.02	−0.03	−0.03	−0.03
Mean Forecast Error	0.47	0.30	0.28	0.29	0.21	0.18	0.18	0.14	0.14
Correlation Between Implied and Realized Rate Changes	0.45	0.32	0.28	0.17	0.04	0.01	−0.01	0.01	0.01

Notes: The Eurodeposit spot yield curves are estimated on the basis of monthly Eurodeposit futures prices between 1987 and 1995. The Treasury spot yield curves are estimated on the basis of OTR bill and bond data. (Note that the price-yield curve of Eurodeposit futures is linear; thus, the convexity bias does not influence the futures-based spot curve. Convexity bias is worth only a couple of basis points for the two-year zeros.)

Interpretations

The empirical evidence in Exhibit 1 is clearly inconsistent with the pure expectations hypothesis. One possible explanation is that curve steepness mainly reflects time-varying risk premiums, and this effect is variable enough to offset the otherwise positive relation between curve steepness and rate expectations. That is, if the market requires a high risk premium, the current long rate will become higher and the curve steeper than what the rate expectations alone would imply — the yield of a long bond initially has to increase so much that it provides the required bond return by its high yield *and* by capital gain caused by its expected rate decline. In this case, rate expectations and risk premiums are negatively related; the steep curve predicts high risk premiums and declining long rates.

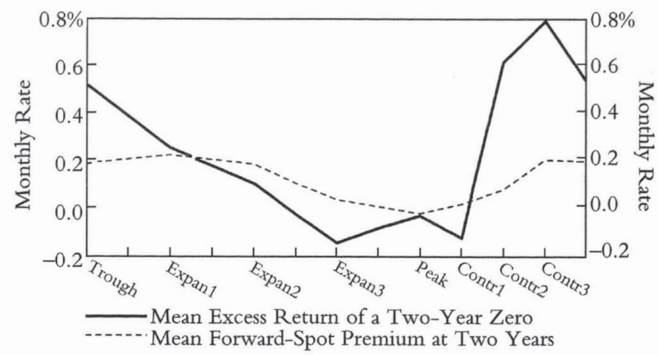
Why should required bond risk premiums vary over time? In general, an asset's risk premium reflects the amount of risk and the market price of risk. Both determinants can fluctuate over time and result in predictability. They may vary with the yield level (rate level-dependent volatility) or market direction (asymmetric volatility or risk aversion) or with economic conditions. For example, cyclical patterns in bond returns may reflect wealth-dependent variation in the risk aversion level.

Exhibit 4 shows the typical business cycle behavior of bond returns and yield curve steepness. Bond returns are high and yield curves are steep near troughs, and bond returns are low and yield curves are flat/inverted near peaks. These counter-cyclic patterns probably reflect the response of monetary policy to the economy's inflation dynamics, as well as time-varying risk premiums (high risk aversion and required risk premiums in "bad times" and vice versa).

Exhibit 4 is constructed so that if bonds tend to earn their rolling yields, the two lines are perfectly aligned. The graph shows that bonds tend to earn additional capital gains (beyond rolling yields) from declining rates near cyclical troughs — and capital losses from rising rates near peaks. Thus, realized bond returns are related to the steepness of the yield curve and — in addition — to the level of economic activity.

These empirical findings motivate the idea that the required bond risk premium varies over time with the steepness of the yield curve and with some other variables. Ilmanen [1997] shows that yield curve steepness indicators and real bond yields, combined

EXHIBIT 4 ■ Average Business Cycle Pattern of U.S. Realized Bond Risk Premium and Curve Steepness, 1968–1995



Notes: Each line is constructed by computing the average value of a series in eight "stage of the business cycle" subsamples. Peak and trough subsamples refer to seven-month windows around the cyclical peaks and troughs, as defined by the National Bureau of Economic Research. In addition, each business cycle is split into three-thirds of expansion and three-thirds of contraction, and each month is assigned to one of these six subsamples. The spacing of subsamples on the x-axis is partially adjusted for the fact that expansions tend to last much longer than contractions. The forward-spot premium measures the steepness of the forward rate curve (the deannualized one-month rate twenty-three months forward minus the current one-month rate). The realized bond risk premium measures the monthly excess return of a synthetic two-year zero-coupon bond over a one-month bill. If the steepness of the forward rate curve is a one-for-one predictor of future excess returns, the two lines are perfectly aligned.

with measures of recent stock and bond market performance, are able to forecast up to 10% of the variation in monthly excess bond returns. For quarterly or annual horizons, the predictable part is even larger. Thus, yield-seeking active strategies appear profitable in the long run.

If market participants are rational, bond return predictability should reflect time variation in the bond risk premiums. Bond returns are predictably high when bonds command exceptionally high risk premiums — either because bonds are particularly risky or because investors are exceptionally risk-averse.

An alternative interpretation is that systematic forecasting errors cause the predictability. If forward rates really reflect the market's rate expectations (and no risk premiums), these expectations are irrational. They tend to be too high when the yield curve is upward sloping and too low when the curve is inverted. The market appears to repeat costly mistakes that it could avoid simply by not trying to forecast rate shifts. Such irrational behavior is inconsistent with market efficiency.

Because expectations are not observable, we can never know to what extent the return predictability reflects time-varying bond risk premiums and systematic forecast errors. Researchers have tried to develop models that explain the predictability as rational variation in required returns. Yield volatility and other obvious risk measures, however, seem to have little ability to predict future bond returns. The observed counter-cyclic patterns in expected returns do suggest rational variation in the risk aversion level — although they also could reflect irrational changes in the market sentiment.

Studies that use survey data to proxy for the market's expectations conclude that risk premiums *and* irrational expectations contribute to the return predictability (see Froot [1989] and De Bondt and Bange [1992]).

II. HOW SHOULD WE INTERPRET THE YIELD CURVE CURVATURE?

The market's curve reshaping expectations, volatility expectations, and expected return structure determine the curvature of the yield curve. Expectations for yield curve flattening imply expected profits for duration-neutral long-barbell versus short-bullet positions, tending to make the yield curve concave (the yield disadvantage of these positions offsets their expected profits from the curve flattening). Expectations for higher volatility increase the value of convexity and the expected profits of these barbell-bullet positions, again inducing a concave yield curve shape. Finally, high required returns of intermediate bonds (bullets) relative to short and long bonds (barbells) make the yield curve more concave. Conversely, expectations for yield curve steepening or for low volatility, together with bullets' low required returns, can even make the yield curve convex.

We analyze the yield curve curvature and focus on two key questions:

1. How important are each of the three determinants in changing the curvature over time?
2. Why is the long-run average shape of the yield curve concave?

Empirical Evidence

Some studies suggest that the curvature of the yield curve is closely related to the market's volatility expectations, presumably because of the convexity bias. Our empirical analysis indicates that the curvature varies more with the market's curve reshaping expectations than with the volatility expectations. The broad curvature of the yield curve varies closely with the steepness of the curve, probably reflecting mean-reverting rate expectations.

Historically, low short rates have been associated with steep yield curves and high curvature (concave shape), while high short rates have been associated with inverted yield curves and negative curvature (convex shape); see Exhibit 5. Steepness measures are negatively correlated with the short rate levels (but almost uncorrelated with the long rate levels), reflecting the higher likelihood of bull steepeners and bear flatteners than bear steepeners and bull flatteners.

The high correlation (0.79) between the changes in the steepness and the changes in the curvature has a nice economic logic. Our curvature measure can be viewed as the yield carry of a curve-steepening position, a duration-weighted bullet-barbell position. If market participants have mean-reverting rate expectations, they expect yield curves to revert to a certain average shape (slightly upward sloping) in the long run. Then, exceptionally steep curves are associated with expectations for subsequent curve flattening and for capital losses on steepening positions. Given the

EXHIBIT 5 ■ Correlation Matrix of Yield Curve Level, Steepness, and Curvature, 1968-1995

	3-Mo. Rate	6-Yr. Rate	Steepness	Curvature
3-Mo. Spot Rate	1.00			
6-Yr. Spot Rate	0.70	1.00		
Steepness	–0.43	–0.04	1.00	
Curvature	–0.20	0.10	0.79	1.00

Notes: The Treasury spot yield curves are estimated on the basis of OTR bill and bond data (see Exhibit 1). The correlations are between the monthly changes in spot rates (or their spreads). Steepness refers to the yield spread between the six-year spot rate and the three-month spot rate. Curvature refers to the yield spread between a synthetic long bullet (three-year zero) and a duration-matched short barbell (0.5 × three-month zero + 0.5 × 5.75-year zero).

expected capital losses, these positions need to offer an initial yield pickup, which leads to a concave (humped) yield curve shape.

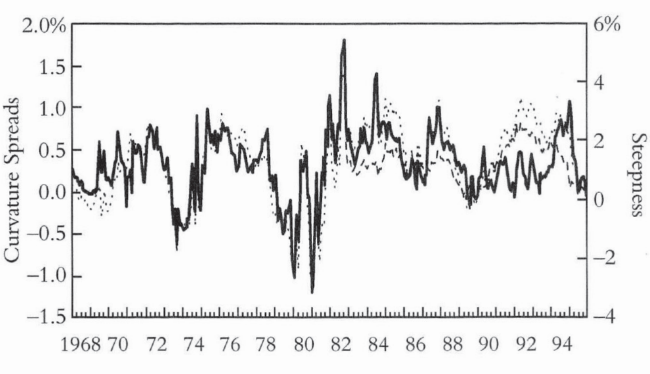
Conversely, abnormally flat or inverted yield curves are associated with the market's expectations for subsequent curve steepening and for capital gains on steepening positions. Given the expected capital gains, these positions can offer an initial yield give-up, which induces a convex (inversely humped) yield curve.

Exhibit 6 illustrates the close comovement between our curve steepness and curvature measures. This pattern is consistent with the mean-reverting rate expectations we have described. Periods of steep yield curves (mid-1980s and early 1990s) are associated with high curvature and, thus, a large yield pickup for steepening positions, presumably to offset their expected losses as the yield curve flattens. Periods of flat or inverted curves (1989-1990 and 1995) are associated with low curvature or even an inverse hump. Thus, barbells can pick up yield and convexity over duration-matched bullets, presumably to offset their expected losses when the yield curve is expected to steepen toward its normal shape.

The expectations for mean-reverting curve steepness influence the broad curvature of the yield curve. In addition, the curvature of the front end sometimes reflects the market's strong view about near-term monetary policy actions and their impact on the curve steepness. Historically, the Federal Reserve Board has tried to smooth interest rate behavior by gradually adjusting the rates that it controls. Such a rate-smoothing policy makes the monetary authority's actions partly predictable and induces a positive autocorrelation in short-term rate behavior. Thus, if the Fed has recently begun to ease (tighten) monetary policy, it is reasonable to expect the monetary easing (tightening) to continue and the curve to steepen (flatten).

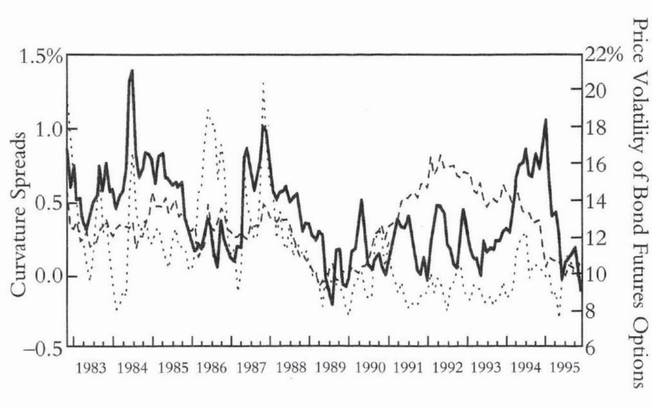
In the literature, yield curve curvature has been mainly associated with the level of volatility. Litterman, Scheinkman, and Weiss [1991] point out that higher volatility should make the yield curve more humped (because of convexity effects) and that a close relation appears to exist between the yield curve curvature and the implied volatility in the Treasury bond futures options. Exhibit 7, however, shows that the relation between curvature and volatility is close only during the sample period of this study (1984-1988). Interestingly, no recessions occurred in the mid-1980s; the yield curve

EXHIBIT 6 ■ Curvature and Steepness of the Treasury Curve, 1968-1995



Notes: Curvature A refers to the yield spread between a long bullet (three-year zero) and a duration-matched short barbell (0.5 × three-month zero + 0.5 × 5.75-year zero). Curvature B refers to the yield spread between a long bullet (ten-year on-the-run bond) and a duration-matched short barbell (duration-weighted combination of two-year and thirty-year on-the-run bonds); this series begins in 1982. Steepness refers to the yield spread between the six-year spot rate and the three-month spot rate.

EXHIBIT 7 ■ Curvature and Volatility in the Treasury Market, 1982-1995



Notes: Curvature A refers to the yield spread between a duration-matched long bullet (three-year zero) and short barbell (0.5 × three-month zero + 0.5 × 5.75-year zero). Curvature B refers to the yield spread between a duration-matched (and cash-neutral) long bullet (ten-year on-the-run bond) and short barbell (duration-weighted combination of two-year and thirty-year on-the-run bonds). Volatility refers to the implied volatility of at-the-money options of the Treasury bond futures; these options began to trade in 1982.

shifts were quite parallel; and the flattening/steepening expectations were probably quite weak.

The relation breaks down before and after the 1984-1988 period — especially near recessions, when the Fed is active and the market may reasonably expect curve reshaping. For example, in 1981 yields were very volatile, but the yield curve was convex (inversely humped). It appears that the market's expectations for future curve reshaping are more important determinants of the yield curve curvature than are its volatility expectations (convexity bias).

The correlations of our curvature measures with the curve steepness are around 0.8, while those with the implied option volatility are around 0.1. Therefore, it is not surprising that the implied volatility estimates that are based on the yield curve curvature are not closely related to the implied volatilities that are based on option prices. Using the yield curve shape to derive implied volatility can result in negative volatility estimates; this unreasonable outcome occurs whenever the expectations for curve steepening make the yield curve inversely humped.

Now we move to the second question: "Why is the long-run average shape of the yield curve concave (steeper at the front end than at the long end)?" This concave shape means that the forwards have, on average, implied yield curve flattening (which would offset the intermediate bonds' initial yield advantage over duration-matched barbells). Exhibit 8 shows that, on average, the implied flattening has not been matched by sufficient realized flattening. Not surprisingly, flattenings and steepenings tend to wash out over time, but the concave spot curve shape is quite persistent.

In fact, a significant positive correlation exists between the implied and the realized curve flattening, but the average forecast errors in Exhibit 8 reveal a bias of too much implied flattening. This conclusion holds when we split the sample into shorter subperiods or into subsamples of a steep versus a flat yield curve environment or a rising rate versus a falling rate environment.

Exhibit 8 shows that, on average, the capital gains caused by the curve flattening have not offset a barbell's yield disadvantage (relative to a duration-matched bullet). A more reasonable possibility is that the barbell's convexity advantage has offset its yield disadvantage. We can evaluate this possibility by examining the impact of convexity on realized returns over time. Empirical evidence suggests that the convexity advantage is not sufficient to offset the yield disadvantage. Alternatively, we can examine the shape of historical average return curve because the realized returns should reflect the convexity advantage.[4]

This convexity effect is certainly a partial explanation for the typical yield curve shape — but it is the sole effect only if duration-matched barbells and bullets have the same expected returns. Equivalently, if the required bond risk premium increases linearly with duration, the average returns of duration-matched barbells and bullets should be the same over a long neutral period (because the barbell's convexity advantage exactly offsets their yield disadvantage).

Historically, the average return curve has been concave, suggesting that bullets have somewhat higher long-run expected returns than duration-matched barbells. Given that the average flattening during the sample is zero, the long-run average concave shape of the yield curve likely reflects the convexity bias and the concave shape of the average bond risk premium curve rather than systematic flattening expectations.

Interpretations

The impacts of curve reshaping expectations and convexity bias on the yield curve shape are easy to understand, but the concave shape of the bond risk pre-

EXHIBIT 8 ■ Evaluating the Implied Forward Yield Curve's Ability to Predict Actual Changes in the Spot Yield Curve's Steepness, 1968-1995

	6 Mo.-3 Mo.	1 Yr.-6 Mo.	3 Yr.-1 Yr.	6 Yr.-3 Yr.	6 Yr.-3 Mo.
Mean Spread (Steepness)	0.33	0.19	0.44	0.32	1.28
Mean Implied Spread Change	–0.33	–0.09	–0.11	–0.05	–0.58
Mean Realized Spread Change	–0.002	–0.001	0.001	0.001	–0.001
Mean Forecast Error	–0.33	–0.09	–0.11	–0.05	–0.58
Correlation Between Implied and Realized Spread Changes	0.53	0.45	0.20	0.03	0.21

mium curve is more puzzling. Why should bullets have a mild expected return advantage over duration-matched barbells? One likely answer is that duration is not the relevant risk measure. We find that average returns are concave even in return volatility, however, suggesting a need for a multifactor risk model.

All one-factor term structure models imply that expected returns should increase linearly with the bond's sensitivity to the risk factor. Because these models assume that bond returns are perfectly correlated, expected returns should increase linearly with return volatility (whatever the risk factor is). Bond durations are proportional to return volatilities only if all bonds have the same basis point yield volatilities, however. Perhaps the concave shape of the average return-duration curve is caused by 1) a linear relation between expected return and return volatility and 2) a concave relation between return volatility and duration that, in turn, reflects an inverted or humped term structure of yield volatility.

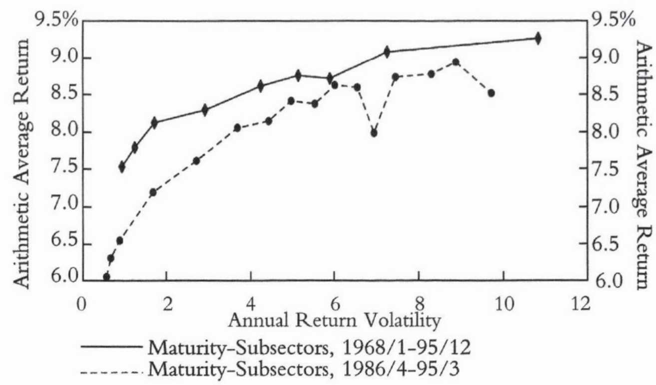
Intuitively, a concave relation between the actual return volatility and duration would make a barbell a more defensive (bearish) position than a duration-matched bullet. The return volatility of a barbell is simply a weighted average of its constituents' return volatilities (given perfect correlation); thus, the barbell's volatility would be lower than that of a duration-matched bullet.

Exhibits 11 and 12 will demonstrate that the empirical term structure of yield volatility has been inverted or humped most of the time. Thus, perhaps a barbell and a bullet with equal return volatilities (as opposed to equal durations) should have the same expected return, but it turns out that the bullet's return advantage persists even when we plot average returns on historical return volatilities.

Exhibit 9 shows the historical average returns of various maturity subsector portfolios of Treasury bonds as a function of return volatility. The average returns are based on two relatively neutral periods, January 1968 to December 1995 and April 1986 to March 1995. We still find that the average return curves have a somewhat concave shape. Note that we demonstrate the concave shape in a conservative way by graphing arithmetic average returns; the geometric average return curves would be even more concave.

One-factor term structure models assume that bond returns are perfectly correlated. One-factor asset pricing models are somewhat more general. They

EXHIBIT 9 ■ Average Maturity Subsector Returns of U.S. Treasuries as a Function of Return Volatility



Notes: The curves show the annualized arithmetic averages of monthly returns of various Treasury bill and bond portfolios. The two curves differ in that we can split the Treasury market into narrower maturity subsector buckets in the more recent sample. The first three points in each curve correspond to constant-maturity three-month, six-month, and nine-month bill portfolios. The next four points correspond to maturity subsector portfolios of 1-2, 2-3, 3-4, and 4-5 year Treasuries. The last two points in the longer sample correspond to a five- to ten-year bond portfolio and a twenty-year bond portfolio. The last nine points in the shorter sample correspond to maturity subsector portfolios of 5-6, 6-7, 7-8, 8-9, 9-10, 10-15, 15-20, 20-25, and 25-30 year Treasuries. Our return calculations ignore the on-the-run bonds' repo market advantage, partly explaining the low returns of the 9-10 year and the 25-30-year Treasury portfolios.

assume that realized bond returns are influenced by only one systematic risk factor but that they also include a bond-specific residual risk component (which can make bond returns imperfectly correlated). Because the bond-specific risk is easily diversifiable, only systematic risk is rewarded in the marketplace. Therefore, expected returns are linear in the systematic part of return volatility.

This distinction is not very important for government bonds because their bond-specific risk is so small. If we plot the average returns on systematic volatility only, the front end would be slightly less steep than in Exhibit 9 because a larger part of short bills' return volatility is asset-specific. Nonetheless, the overall shape of the average return curve would remain concave.

Convexity bias and the term structure of yield volatility partly explain the concave shape of the average yield curve, but a non-linear expected return curve appears to be an additional reason. Exhibit 9 suggests that expected returns are somewhat concave in return volatility. That is, long bonds have lower required returns than one-factor models imply. Some desirable property in the longer cash flows makes the market accept a lower expected excess return per unit of return volatility for them than for the intermediate cash flows.

We need a second risk factor, besides the rate level risk, to explain this pattern. The pattern moreover may teach us something about the nature of the second factor and about the likely sign of its risk premium.

Volatility as the second factor could explain the observed patterns if the market participants, in the aggregate, prefer insurance-type or "long-volatility" payoffs. Even non-optionable government bonds have an option-like characteristic because of the convex shape of their price-yield curves. The value of convexity increases with a bond's convexity and with the perceived level of yield volatility. If volatility does not influence expected returns, a yield disadvantage exactly offsets longer bonds' convexity advantage. The concave shape of the average return curve in Exhibit 9, however, suggests that positions that benefit from higher volatility have lower expected returns than positions that are adversely affected by higher volatility. Findings from the options market and the mortgage bond market are consistent with this idea (see Goodman and Ho [1997]).

Although the evidence is tentative, we find the negative sign for the price of volatility risk intuitively appealing. The Treasury market participants may be especially averse to losses in high-volatility states, or they may prefer insurance-type (skewed) payoffs so much that they accept lower long-run returns for them. Thus, the long bonds' low expected return could reflect the high value many investors assign to positive convexity. Yet because short bonds exhibit little convexity, other factors are needed to explain the curvature at the front end of the yield curve.

Yield curve steepness as the second factor (or short rate and long rate as the two factors) could explain the observed patterns if curve-flattening positions tend to be profitable just when investors value profits most. We do not think that the curve steepness is by itself a risk factor that investors worry about, but it may tend to coincide with a more fundamental factor. Recall that the concave average return curve suggests that self-financed curve-flattening positions have negative expected returns — because they are more sensitive to the long rates (with low expected returns per volatility) than to the short/intermediate rates (with high expected returns per volatility).

This negative risk premium has a theoretical justification if the flattening trades are especially good hedges against "bad times." An academic's description of bad times would be a period of high marginal utility of profits; a practitioner's definition probably is a deep recession or a bear market.

The empirical evidence on this issue is mixed. It is clear that long bonds performed very well in deflationary recessions (the United States in the 1930s, Japan in the 1990s). They did not perform at all well in the stagflations of the 1970s when the predictable and realized excess bond returns were negative. Since World War II, U.S. long bond performance has been positively correlated with stock market performance — although bonds turned out to be a good hedge during the stock market crash of October 1987.

Flattening positions have not been good recession hedges either; the yield curves typically have been flat or inverted at the beginning of a recession and have steepened during it (see also Exhibit 4). Nonetheless, flattening positions typically have been profitable in a rising rate environment; thus, they have been reasonable hedges against a bear market *for bonds*.

III. HOW DOES THE YIELD CURVE EVOLVE OVER TIME?

Many popular term structure models allow the decomposition of forward rates into a rate expectation component, a risk premium component, and a convexity bias component, but they make different assumptions about the behavior of the yield curve over time. Specifically, the models differ in their assumptions regarding the number and identity of factors influencing interest rates, the factors' expected behavior (the degree of mean reversion in short rates and the role of a risk premium), and the factors' unexpected behavior (for example, the dependence of yield volatility on the yield level). These characteristics of yield curve behavior are relevant for evaluating the realism of various term structure models.[5]

The simple model of only parallel shifts in the spot curve makes extremely restrictive and unreasonable assumptions — for example, it does not preclude negative interest rates.[6] In fact, it is equivalent to the

Vasicek [1977] model with no mean reversion. All one-factor models imply that rate changes are perfectly correlated across bonds. The parallel shift assumption requires, in addition, that the basis point yield volatilities are equal across bonds.

Other one-factor models may imply other (deterministic) relations between the yield changes across the curve, such as multiplicative shifts or greater volatility of short rates than of long rates. Multifactor models are needed to explain the observed imperfect correlations across bonds — as well as the non-linear shape of expected bond returns as a function of return volatility that is discussed above.

Time Series Evidence

In our brief survey of empirical evidence, it is useful to focus first on the time series implications of various models and then on their cross-sectional implications. We begin by examining the expected part of yield changes, or the degree of mean reversion in interest rate levels and spreads.

If interest rates follow a random walk, the current interest rate is the best forecast for future rates — that is, changes in rates are unpredictable. In this case, the correlation of (say) a monthly change in a rate with the beginning-of-month rate level or with the previous month's rate change should be zero. If interest rates do not follow a random walk, these correlations need not equal zero. In particular, if rates are mean-reverting, the slope coefficient in a regression of rate changes on rate levels should be negative. That is, falling rates should follow abnormally high rates, and rising rates should succeed abnormally low rates.

Exhibit 10 shows that interest rates do not exhibit much mean reversion. The slope coefficients of yield changes on yield levels are negative, consistent with mean reversion, but they are not quite statistically significant. Yield curve steepness measures are more mean-reverting than yield levels. Mean reversion is more apparent at the annual horizon than at the monthly horizon, consistent with the idea that mean reversion is slow. In fact, yield changes seem to exhibit some trending tendency in the short run (the autocorrelations between the monthly yield changes are positive), until a "rubber-band effect" begins to pull yields back when they get too far from the perceived long-run mean.

If we focus on the evidence from the 1990s (not shown), the main results are similar to those in Exhibit 10, but short rates are more predictable (more mean-reverting and more highly autocorrelated) than long rates, probably reflecting the Fed's rate-smoothing behavior.

Moving to the unexpected part of yield changes, we analyze the behavior of (basis point) yield volatility over time. In an influential study, Chan, Karolyi, Longstaff, and Sanders [1992] show that various specifications of common one-factor term structure models differ in two respects: the degree of mean reversion and the level dependence of yield volatility. Empirically, they find insignificant mean reversion and significantly level-dependent volatility — more than a one-for-one relation.[7]

EXHIBIT 10 ■ Mean Reversion and Autocorrelation of U.S. Yield Levels and Curve Steepness, 1968-1995

	3 Mo.		2 Yr.		30 Yr.		2 Yr.-3 Mo.		30 Yr.-2 Yr.	
	Mon	Ann	Mon	Ann	Mon	Ann	Mon	Ann	Mon	Ann
Mean Reversion Coeff.	−0.03	−0.26	−0.02	−0.26	−0.01	−0.20	−0.13	−0.74	−0.06	−0.42
t-statistic	−1.26	−1.75	−1.20	−1.74	−1.10	−1.36	−3.40	−2.49	−2.19	−2.70
R^2	2%	12%	1%	13%	1%	10%	7%	37%	3%	22%
First Autocorrel. Coeff.	0.10	0.24	0.17	−0.06	0.15	−0.10	−0.12	−0.21	0.08	0.10
t-statistic	0.87	1.26	2.46	−0.39	2.08	−0.46	−1.39	−1.48	1.08	0.56
R^2	1%	6%	3%	1%	2%	1%	1%	4%	1%	1%

Notes: Numbers are based on the on-the-run yields of a three-month bill, a two-year note, and a thirty-year (or the longest available) bond. We use 335 monthly observations and 27 annual observations. The mean reversion coefficient is the slope coefficient in a regression of each yield change on its beginning-of-period level. The first-order autocorrelation coefficient is the slope coefficient in a regression of each yield change on the previous period's yield change. The (robust) t-statistics measure the statistical significance, and the (unadjusted) R^2 values measure the explanatory power of the regression.

Moreover, they find that the evaluation of various one-factor models' realism depends crucially on the volatility assumption; models that best fit U.S. data have a level-sensitivity coefficient of 1.5. According to these models, future yield volatility depends on the current rate level and nothing else: High yields predict high volatility.

Another class of models — so-called GARCH models — stipulate that future yield volatility depends on the past volatility; high recent volatility and large recent shocks (squared yield changes) predict high volatility. Brenner, Harjes, and Kroner [1996] show that empirically the most successful models assume that yield volatility depends on the yield level *and* on past volatility. With GARCH effects, the power coefficient drops to approximately 0.5.

Finally, all these studies include the exceptional period 1979-1982, which dominates the results (see Exhibit 11). In this period, yields rose to unprecedented levels — but the increase in yield volatility was even more extraordinary. Since 1983, the U.S. yield volatility has varied much less closely with the rate level.[8]

Cross-Sectional Evidence

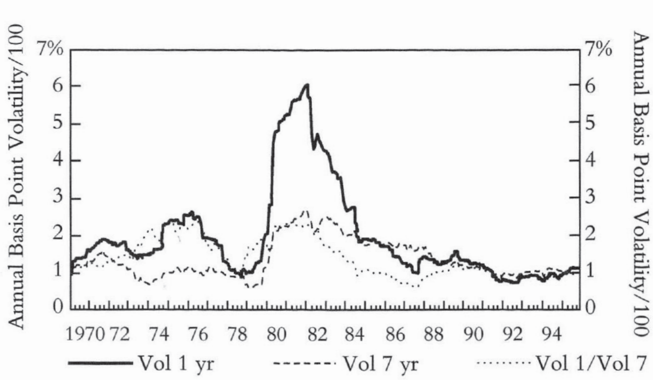
We first discuss the shape of the term structure of yield volatilities and its implications for bond risk measures and later describe the correlations across various parts of the yield curve. The term structure of basis point yield volatilities in Exhibit 12 is steeply inverted when we use a long historical sample period. Theoretical models suggest that the inversion in the volatility structure is mainly due to mean-reverting rate expectations. Intuitively, if long rates are perceived as averages of expected future short rates, extreme variation in the short rates would have a lesser impact on the long rates.

The observation that the term structure of volatility inverts quite slowly is consistent with expectations for very slow mean reversion. In fact, since the 1979-1982 period, the term structure of volatility has been reasonably flat — as evidenced by the ratio of short rate volatility to long rate volatility in Exhibit 11.

The subperiod evidence in Exhibit 12 confirms that the term structure of volatility has recently been humped rather than inverted. The upward slope at the front end of the volatility structure may reflect the Fed's smoothing (anchoring) of very short rates while the one- to three-year rates vary more freely with the market's rate expectations and with the changing bond risk premiums.

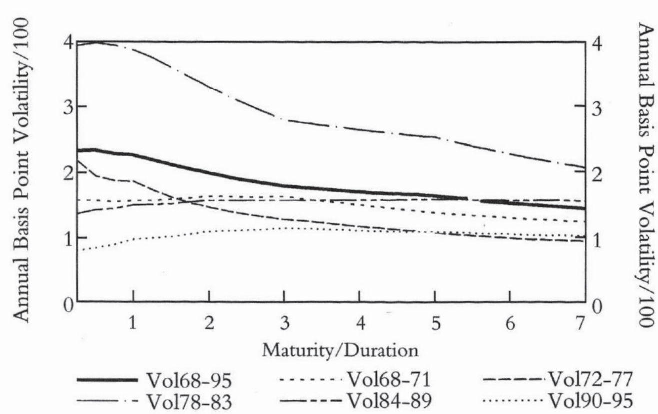
The non-flat shape of the term structure of yield volatility has important implications for the relative riskiness of various bond positions. The traditional duration is an appropriate risk measure only if the yield volatility structure is flat. We pointed out earlier that inverted or humped yield volatility structures would make the

EXHIBIT 11 ■ 24-Month Rolling Spot Rate Volatilities



Notes: The graphs plot the annualized volatilities of the monthly basis point changes in one-year and seven-year spot rates (and their ratio) over twenty-four-month subperiods.

EXHIBIT 12 ■ Term Structure of Spot Rate Volatilities



Notes: Each volatility term structure plots the annualized volatilities of the monthly basis point changes in various-maturity spot rates over a given sample period. For example, the three-month rate's volatility was 2.32% (232 basis points) for the full sample

EXHIBIT 13A ■ Basis Point Yield Volatility for Various Model Specifications

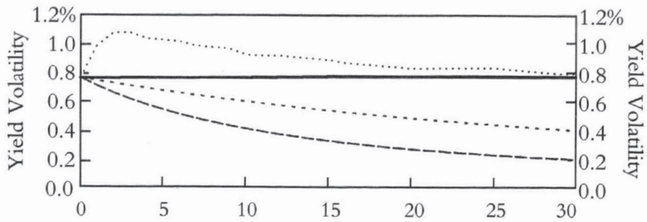
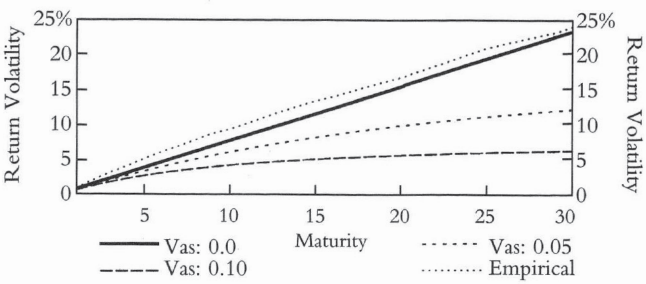


EXHIBIT 13B ■ Return Volatility for Various Model Specifications



Notes: Four term structures of annualized basis point spot yield volatilities and the corresponding return volatilities. The empirical yield volatilities are based on weekly Treasury market data between 1990-1995, while the three other volatility structures are based on varying specifications of the Vasicek model — one with no mean reversion (the world of parallel shifts), another with mild mean reversion ($k = 0.05$ in $dr = k(l - r)dt + sdz$), and a third one with stronger mean reversion ($k = 0.10$). The return volatilities are computed for each zero-coupon bond by the product of basis point yield volatility and duration. These return volatilities are proportional to the factor sensitivities in one-factor models.

return volatility curve a concave function of duration.

Exhibit 13 shows examples of flat, humped, and inverted yield volatility structures — and the corresponding return volatility structures. The humped volatility structure reflects empirical yield volatilities in the 1990s, while the flat and inverted volatility structures are based on the Vasicek model with mean reversion coefficients of 0.00, 0.05, and 0.10. The model's short rate volatility is calibrated to match that of the three-month rate in the 1990s (77 bp or 0.77%).

It is clear from this graph that the traditional duration exaggerates the relative riskiness of long bonds whenever the term structure of yield volatility is inverted or humped. Moreover, the relative riskiness will be quite misleading if the assumed volatility structure is inverted (as in the long sample period in Exhibit 12) while the actual volatility structure is flat or humped (as in the 1990s).

Historical analysis shows that correlations of yield changes across the Treasury yield curve are not perfect but are typically very high beyond the money market sector (0.82-0.98 for the monthly changes of the two- to thirty-year on-the-run bonds between 1968-1995) and reasonably high even for the three-month bills and thirty-year bonds (0.57).

Thus, the evidence is not consistent with a one-factor model. Rather, it appears that two or three statistically derived systematic factors can explain 95%-99% of the fluctuations in the yield curve (see Garbade [1986] and Litterman and Scheinkman [1991]). The patterns of sensitivities to each factor across bonds of different maturities indicate that the three most important factors can be interpreted as the level, slope, and curvature factors.

ENDNOTES

This article is an abbreviated version of a Salomon Brothers research report titled "The Dynamics of the Shape of the Yield Curve — Understanding the Yield Curve: Part 7." It was written while both authors were working at Salomon Brothers Inc, New York. The authors thank Nick Clarke, Robert McAdie, Richard Pagan, David Pollard, Janet Showers, and Charlie Ye for helpful comments and Bin Shao for his technical contributions.

[1]Another way to get around the problem that the market's rate expectations are unobservable is to assume that a survey consensus view can proxy for these expectations. A comparison of the forward rates with survey-based rate expectations indicates that changing rate expectations and changing bond risk premiums induce changes in the curve steepness (see Exhibit 9 in Ilmanen [1996]).

[2]The deviations from the pure expectations hypothesis *are* statistically significant when we regress excess bond returns on the steepness of the forward rate curve. Moreover, as long as the correlations in Exhibit 1 are zero or below, long bonds tend to earn *at least* their rolling yields.

[3]Ilmanen [1996] shows that the forwards have predicted future excess bond returns better than they have anticipated future yield changes. Ilmanen [1997] discusses more general evidence of the forecastability of excess bond returns. In particular, combining yield curve information with other predictors can enhance the forecasts.

[4]See Figures 11 and 12 in Ilmanen [1995].

[5]We provide empirical evidence on the historical behavior of nominal interest rates. This evidence is not directly relevant for evaluating term structure models in some important situations. First, when term structure models are used to value derivatives in an arbitrage-free framework, these models make assumptions concerning the risk-neutral probability distribution of interest rates, not concerning the real-world distribution. Second, equilibrium term structure models often describe the behavior of real interest rates, not nominal rates.

[6]Moreover, a model with parallel shifts would offer riskless arbitrage opportunities if the yield curves were flat. Duration-matched long barbell versus short bullet positions with positive convexity could either be profitable or break even because there would be no yield give-up or any possibility of capital losses caused by the curve steepening. The parallel shift model would not offer riskless arbitrage opportunities if the spot curves were concave (humped) because the barbell-bullet positions' yield give-up could more than offset their convexity advantage.

[7]In many term structure models, short rate volatility can be expressed as proportional to r^γ where γ is the power coefficient of volatility's sensitivity on the rate level; see Iwanowski [1996]. For example, in the Vasicek [1977] model (additive or normal rate process), $\gamma = 0$, while in the Cox, Ingersoll, and Ross [1985] model (square root process), $\gamma = 0.5$. The Black, Derman, and Toy [1990] model (multiplicative or lognormal rate process) is not directly comparable but $\gamma \approx 1$. Intuitively, if $\gamma = 0$, the basis point yield volatility [$\mathrm{Vol}(\Delta y)$] does not vary with the yield level, while if $\gamma = 1$, the basis point yield volatility varies one-for-one with the yield level — and the relative yield volatility [$\mathrm{Vol}(\Delta y/y)$] is independent of the yield level.

[8]When we estimate the coefficient (short rate volatility's sensitivity to the rate level) using daily changes of the three-month Treasury bill rate, we find that the coefficient falls from 1.44 between 1977-1994 to 0.71 between 1983-1994. Moreover, when we reestimate the coefficient in a model that accounts for simple GARCH effects, it falls to 0.37 and 0.17, suggesting little level dependence. (The GARCH coefficient on the past variance is 0.87 and 0.95 in the two samples, and the GARCH coefficient on the previous squared yield change is 0.02 and 0.03.)

REFERENCES

Black, F., E. Derman, and W. Toy. "A One-Factor Model of Interest Rates and Its Application To Treasury Bond Options." *Financial Analysts Journal*, 46, 1 (1990).

Brenner, R., R. Harjes, and K. Kroner. "Another Look at Models of the Short-Term Interest Rate." *Journal of Financial and Quantitative Analysis*, 31, 1 (1996).

Chan, K., G.A. Karolyi, F. Longstaff, and A. Sanders. "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate." *Journal of Finance*, 47, 3 (1992).

Cox, J.C., J.E. Ingersoll, and S.A. Ross. "A Theory of the Term Structure of Interest Rates." *Econometrica*, 53, 2 (1985).

De Bondt, W.F.M., and M.M. Bange. "Inflation Forecast Errors and Time Variation in Term Premia." *Journal of Financial and Quantitative Analysis*, 27, 4 (1992).

Froot, K. "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates." *Journal of Finance*, 44, 2 (1989).

Garbade, K. "Modes of Fluctuations in Bond Yields — An Analysis of Principal Components." Bankers Trust Co., 1986.

Goodman, L., and J. Ho. "Are Investors Rewarded for Shorting Volatility?" *Journal of Fixed Income*, June 1997.

Ilmanen, A. "Convexity Bias and the Yield Curve. Understanding the Yield Curve: Part 5." Salomon Brothers Inc, 1995.

——. "Forecasting U.S. Bond Returns." *Journal of Fixed Income*, June 1997.

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

Iwanowski, R. "A Survey and Comparison of Term Structure Models." Unpublished manuscript, Salomon Brothers Inc, 1996.

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

Litterman, R., J. Scheinkman, and L. Weiss. "Volatility and the Yield Curve." *Journal of Fixed Income*, June 1991.

Vasicek, O. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics*, 5, 2 (1977).