

The Journal of
FIXED INCOME

VOLUME 7 NUMBER 1

www.ijjfi.com

JUNE 1997

FORECASTING U.S. BOND RETURNS

ANTTI ILMANEN

IIJ Institutional
Investor
Journals

The Voices of Influence | ijjfi.com

FORECASTING U.S. BOND RETURNS

ANTTI ILMANEN

It is difficult to forecast bond market fluctuations, although some research shows that these fluctuations are not fully unpredictable: It is possible to identify in advance periods when the reward for duration extension is likely to be abnormally high or abnormally low.

In this article, we first describe a few variables that have the ability to predict near-term bond market performance. We then show how to combine the information that these predictors provide into a single forecast and, further, into implementable investment strategies. Finally, we backtest the historical performance of these strategies in a realistic out-of-sample setting.

We show elsewhere that intermediate- and long-term bonds earn higher average returns than short-term bonds (Ilmanen [1996a]). This evidence suggests that the long-run bond risk premium is positive.¹ If the risk premium is constant over time, the long-run average risk premium is also our best forecast for the near-term bond market performance.

Steeply upward-sloping yield curves tend to precede high excess bond returns, and inverted yield curves tend to precede negative excess bond returns. It follows then that the risk premium is not constant, and that the current shape of the yield curve provides valuable information about the time-varying bond risk premium. In this article, we show that other variables can enhance the yield curve's ability to forecast near-term excess bond returns.

The predictability of excess bond returns has important implications for investors who are willing to use so-called tactical asset allocation strategies. According to our extensive historical analysis, strategies

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

that adjust portfolio duration dynamically using signals from the predictor variables would have earned substantially higher long-run returns than did static strategies that do not actively adjust the portfolio duration.

I. REASON FOR A CHANGING RISK PREMIUM: A CYCLE OF FEAR AND GREED

Empirical research shows that the expected excess returns of stocks and long-term bonds vary over the business cycle; they tend to be high at the end of recession and low at the end of expansion. Are there any intuitive reasons that expected excess returns should vary with economic conditions?

If the expected returns reflect rational risk premiums, they change over time as the amount of risk in assets or the market price of risk vary over time. The risk premium that bond market participants expect should increase if the risk increases (because of higher interest rate volatility, higher covariance between bonds and stocks, or greater inflation uncertainty). In addition, the required risk premium should increase if the market participants' aggregate risk aversion level increases.

We propose that investors are more risk-averse (fearful) when their current wealth is low relative to their past wealth. Greater risk aversion makes them demand higher risk premiums (more compensation for holding risky assets) near business cycle troughs. Conversely, greater wealth near business cycle peaks makes investors less risk-averse (more complacent, or greedy); therefore, they bid down the required risk premium (by bidding up the asset prices).

Thus, the observed time variation in risk premiums may be explained by the Wall Street adage about the cycle of fear and greed. Authors who discuss wealth-dependent changes in risk aversion level include Marcus [1989], Sharpe [1990], Chen [1991], and Campbell and Cochrane [1995].

II. ARE EXCESS BOND RETURNS PREDICTABLE?

Which Variables Forecast Excess Bond Returns?

As mentioned above, measures of yield curve steepness have some ability to predict subsequent excess

bond returns. Elsewhere we show that a steep yield curve may reflect a high required risk premium or the market's expectations of rising rates (Ilmanen [1996b]).

If the expectation is assumed to be zero (the current yield curve is the market's best forecast for future yield curves), then yield curve steepness is a good proxy for the bond risk premium. We measure the curve steepness by the term spread (the difference between a long-term rate and a short-term rate).

Conveniently, we can use the term spread as an overall proxy for the bond risk premium, even if we do not know what causes the expected return differentials across bonds. For this reason, the term spread is our first predictor variable.

Yet, if we are trying to forecast bond returns, why restrict ourselves to just one predictor? It is likely that the term spread is sometimes influenced by the market's rate expectations, which makes the zero-change assumption unrealistic. Because rate expectations are unobservable, we cannot know how much "noise" they introduce in our risk premium proxy.

Thus, we do not know to what extent a given shape of the curve reflects the required bond risk premium and to what extent it reflects the market's rate expectations. Using other predictor variables to with the term spread should help us *filter out* the noise and give us a better signal about the future risk premium.

The filter variables should be correlated with the risk premium. Given our hypothesis of wealth-dependent risk aversion, we combine the information in the term spread and in the stock market's recent performance. The *inverse* of the recent stock market performance is our proxy for the (unobservable) aggregate level of risk aversion.

If a high term spread coincides with a depressed stock market, the curve steepness is less likely to reflect rising rate expectations (because monetary policy tightening and inflation threat are less likely in this environment), and more likely to reflect high required risk premiums (because low stock prices may reflect high required returns on risky assets — or even cause them via wealth-dependent risk aversion). We measure recent stock market performance by "inverse wealth," a weighted average of past returns, where more distant observations have lower weights.

As a third predictor, we will examine the real bond yield, which is sometimes used as the overall proxy for the bond risk premium instead of the term spread. This measure incorporates inflation rate into the

forecasting model.

Our final predictor, momentum, is a dummy variable that simulates a simple moving-average trading rule to exploit the persistence (positive autocorrelation) in bond returns.² This strategy tries to capture large trending moves in the bond market.

To reduce trading when the market is oscillating within a narrow trading range, and thereby avoid “whipsaw” losses from buying at low yields and selling at high yields, we impose a neutral trading range in which no position is held. Somewhat arbitrarily, we use a six-month moving-average window and a 10-basis point neutral trading range.

Thus, the rule is to take a long (short) position in the bond market when the long-term bond yield declines (increases) to more than 5 basis points below (above) its six-month moving average; such a break-out from a trading range is attributed to positive (negative) momentum in the bond market. If the bond yield returns to (stays within) the 10-basis point range around its six-month average, the rule is to take (retain) a neutral position. The dummy variable takes values 1, -1, or 0 if the strategy is long, short, or neutral.

Our empirical analysis will confirm that the term spread can forecast future excess bond returns, but combining the information contained in several predictor variables improves these forecasts further. Linear regression is the most common way to combine the information in several variables.³

We run a multiple regression of the realized excess bond return on the term spread, the real bond yield, and measures of recent stock and bond market performance, and we use the fitted value from this regression as an estimate of the (expected) bond risk premium.

Correlations Between Predictor Variables and Subsequent Excess Bond Returns

We examine the predictability of the monthly excess return of a twenty-year Treasury bond over the one-month bill rate between January 1965 and July 1995. We focus on four predictor variables as described in Exhibit 1: the term spread; the real bond yield; inverse wealth; and momentum.⁴

Exhibit 2 shows the correlations between the excess bond return and various predictor variables.⁵

EXHIBIT 1 ■ Description of the Predictor Variables and the Predicted Variable

Variable	Definition	Data Source
Term Spread	Difference between the estimated five-year spot rate and the three-month spot rate	Center for Research in Security Prices at the University of Chicago and Salomon Brothers since 1994
Real Yield	Difference between the estimated five-year spot rate and the most recently published yearly consumer price inflation rate	Center for Research in Security Prices at the University of Chicago and Salomon Brothers since 1994
Inverse Wealth	Ratio of the exponentially weighted past stock market level to the current stock market level (W_t). Formally = $(W_{t-1} + 0.9 W_{t-2} + 0.9^2 W_{t-3} + \dots +) \times 0.1/W_t$	Ibbotson Associates — Standard & Poor's 500 total return index
Momentum	Dummy variable that takes value 1 if the bond yield is more than 5 basis points below its six-month average, -1 if the bond yield is more than five basis points above its six-month average, and 0 otherwise	Ibbotson Associates — yield of a long-term government bond with an approximate maturity of 20 years
Excess Bond Return	Monthly return of a long-term Treasury bond in excess of the nominally riskless return of a one-month Treasury bill. Also called realized bond risk premium	Ibbotson Associates — total return index of a long-term government bond with an approximate maturity of 20 years

Note: All rates and returns are compounded continuously.

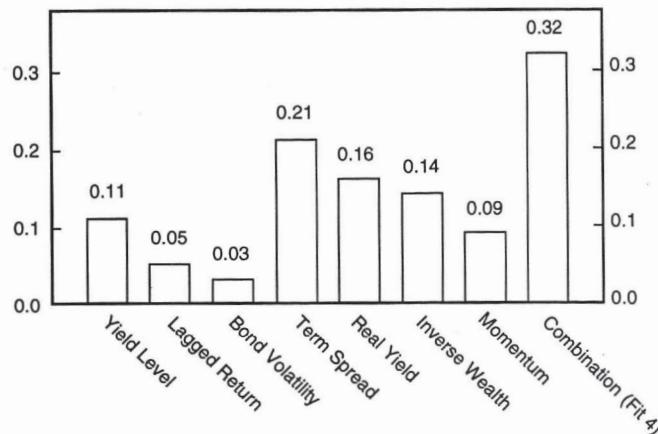
The conventional view that risk premiums cannot be forecast using available information implies that all these correlations should be very close to zero. This conventional view is based partly on the finding that some obvious predictor candidates have limited forecasting ability.

For example, the first three bars in Exhibit 2 show that a bond's yield level, its lagged monthly return, and its past volatility (measured by the twelve-month rolling standard deviation of monthly excess returns) all have low correlations with next month's excess bond return (0.03–0.11). In contrast, the predictors that we have identified above — the term spread, the real yield, inverse wealth, and momentum — have correlations with the subsequent excess bond return between 0.09 and 0.21.

Note that our momentum variable, which is based on a moving-average strategy, has somewhat better forecasting ability than a simple lagged return (which could be used as an alternative proxy for the market's momentum). Finally, combining the information in these four predictors gives even more accurate return predictions, with a correlation of 0.32.⁶ Steep yield curves, high real yields, depressed stock markets, and rallying bond markets are all positive indicators of subsequent bond market performance.

Our predictor variables are financial market data. Many bond market participants are more accustomed to forecasting market movements on the available "fundamental" macroeconomic data. Previous empirical

EXHIBIT 2 ■ Correlation of Various Predictors with Subsequent Monthly Excess Bond Return, 1965–1995



research suggests, however, that *financial market variables* are better predictors of asset returns than *macroeconomic variables* such as production growth rates, perhaps because the latter are less accurately measured and less timely. While market-based variables are forward-looking (partly reflecting the market's expectations about future economic developments), contemporaneous macroeconomic data describe past events, and with a publication lag.

Another finding worth noting from previous studies is the low correlation between various risk measures (such as volatility in Exhibit 2) and future bond returns; periods of high risk do not seem to provide bondholders with high near-term expected returns.⁷

Correlations are not the only way to show that our predictors can discriminate between good and bad times to hold long-term bonds. In Exhibit 3, we examine the average monthly returns in subsamples that are based on the beginning-of-month values of the term spread and inverse wealth.

The annualized average excess return is -12.4% in months that begin with an inverted curve and 2.6% in months that begin with an upward-sloping curve (87% of the time). This finding is consistent with the hypothesis of wealth-dependent risk aversion. Periods of steep yield curves and high risk premiums tend to coincide with cyclical troughs (high risk aversion), while periods of flat or inverted yield curves and low risk premiums tend to coincide with cyclical peaks (low risk aversion). Future bond returns also tend to be higher when inverse wealth is high (the stock market is depressed) than when it is low.

Combining the information in these two predictors sharpens our return predictions further. The average excess bond return is higher when an upward-sloping yield curve coincides with a depressed stock market (12%) than when it coincides with a strong stock market (1%). In the latter case, the curve steepness is more likely to reflect the market's expectations about rising rates than about the required bond risk premium.

The inverse relation between stock market level and subsequent bond returns may be interpreted in many ways. We proposed earlier that declining wealth level makes investors more risk-averse and increases the risk premium they require for holding risky assets.

Alternatively, the relation may be caused by lagged portfolio flows. Poor recent stock market performance can make investors shift money to bonds,

EXHIBIT 3 ■ Excess Bond Return in Subsamples, 1965-1995

Months Beginning With	Term Spread > 0	Term Spread < 0		
Average	2.63%	-12.40%		
No. of Months of Total	87	13		
Months Beginning With	Inverse Wealth > 1	Inv. Wealth < 1		
Average	6.24%	-0.36%		
No. of Months of Total	16	84		
Month Beginning With	Term Sp. > 0 & Inv. Wealth > 1	Term Sp. > 0 & Inv. Wealth < 1	Term Sp. < 0 & Inv. Wealth > 1	Term Sp. < 0 & Inv. Wealth < 1
Average	11.95%	1.23%	-9.76%	-13.60%
No. of Months of Total	12	75	4	9

Note: Average is the annualized average of a twenty-year bond's monthly excess return in each subsample, expressed in percent.

either because these are less risky, or because investors extrapolate and expect the poor stock market performance to continue.

More generally, the time variation in expected returns may reflect rational factors (time-varying risk or risk aversion level) or irrational factors (such as swings in market sentiment or an underreaction of long-term rate expectations to current inflation shocks).

Another way to think about the patterns in the next-to-last row in Exhibit 3 is that when *both* bond and stock markets appear to be "cheap" (the term spread is high, and inverse wealth is high), investors can rely more on these cheapness indicators and expect high future returns for risky assets. Conversely, when *both* bond and stock market indicators signal "richness" (the term spread is negative, and inverse wealth is low), investors can more confidently expect low future returns.

In this light, our three first predictors may be viewed as "value" indicators that tend to give buy signals when asset markets are weak. These predictors are complemented by the fourth one — momentum —

EXHIBIT 4 ■ Results from Regressing Excess Bond Return on Four Predictors, 1965-1995

	Coefficient	T-Statistic
Constant	-10.15	-4.98
Term Spread	0.37	2.24
Real Yield	0.20	2.89
Inverse Wealth	9.85	4.81
Momentum	0.34	2.02
R ²		10.3%

which gives a buy signal when bond prices are trending higher.

Exhibit 4 shows the regression results for our whole sample period. All four predictors are statistically significantly related to subsequent excess bond returns. The regression coefficients show that the expected excess bond returns are high when the yield curve is steeply upward-sloping, the real yield is high, the stock market is depressed, and the bond market has positive momentum.

Together, the four predictors capture 10% of monthly variation in excess bond returns. The fact that 90% of the return variation is unpredictable tells us that even if strategies that exploit these patterns are profitable, they certainly will not be riskless.

Out-of-Sample Estimation of Return Predictions

The regression splits each month's excess bond return into a fitted part and a residual. The fitted part can be viewed as the expected excess bond return, the residual as the unexpected excess bond return. Because the current value of each predictor is known, we can compute the current forecast for the near-term excess bond return by using the equation:

$$\text{Expected Excess Bond Return} = -10.15 + 0.37 \times \text{Term Spread} + 0.20 \times \text{Real Yield} + 9.85 \times \text{Inverse Wealth} + 0.34 \times \text{Momentum}$$

We are using only available information to make this forecast; we combine current values of the predic-

tors with historical estimates of the regression coefficients. For this reason, we can call this an *out-of-sample* forecast, as opposed to an *in-sample* forecast.

For an *in-sample* forecast, we could combine the predictor values at the beginning of January 1965 with the regression coefficients and treat the fitted value as the expected excess bond return for January 1965. In doing so, we would be peeking into the future and assuming that investors already knew the regression coefficients in 1965.

In reality, of course, these coefficients are estimated in 1995 using data from 1965 to 1995. The use of *in-sample* forecasts adds an element of hindsight to the analysis, which leads, at best, to an exaggerated view of return predictability and, at worst, to totally spurious findings. Many investors find the use of *in-sample* forecasts unrealistic.

In general, investors are well-advised to suspect a data-snooping bias when faced with any "exciting" new empirical findings. Data snooping describes the search for profitable regularities in financial market data. The bias arises because some apparently significant findings are likely to be period-specific and spurious.

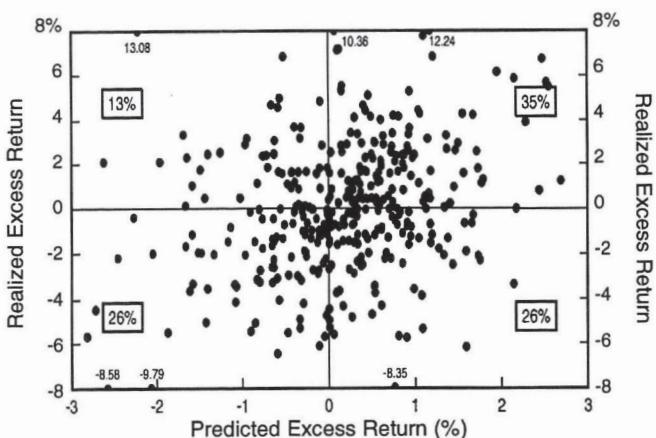
We try to guard against a data-snooping bias by conducting *out-of-sample* analysis, making return predictions each month using only data that are available at the time of forecasting. When we make the forecast of the monthly U.S. excess bond return for January 1965, we run a regression using all historical data from January 1955 to December 1964. The forecast (the fitted part of the regression) combines the estimated regression coefficients with the values of the predictors at the end of December 1964.

To make the forecast for February 1965, we run another regression that uses data from January 1955 to January 1965. We run these monthly rolling regressions with an expanding historical sample until July 1995. This process gives us a series of monthly *out-of-sample* excess bond return forecasts.⁸

We begin the evaluation of the *out-of-sample* forecasts' predictive ability by showing in Exhibit 5 a scatter plot of realized monthly excess bond returns on the *out-of-sample* predictions. To enhance visual clarity, we trim the range of the y-axis to (-8%, +8%) and identify six exceptional observations on the borders.

If the forecasts tend to have correct signs (realized excess returns are positive when they were predicted to be positive and negative when they were predicted to be negative), most observations will lie in the

EXHIBIT 5 ■ Realized Excess Return versus Predicted Excess Return, 1965-1995



upper-right quadrant or in the lower-left quadrant. Without any forecasting ability, all quadrants should contain 25% of the observations.

Exhibit 5 shows that the forecasts have the correct sign in 61% ($= 35 + 26$) of the months. These odds are better than fifty-fifty, but clearly the forecasts are not infallible.

The relation between the predicted and the realized excess returns in Exhibit 5 may not appear very impressive, reflecting the fact that most of the short-term fluctuations in excess bond returns are unpredictable. Perhaps the long-term fluctuations are more predictable; averaging many monthly returns will smooth the return series and may increase the share of the predictable returns.

Some unpublished analysis shows that in a scatter plot of the subsequent twelve-month realized excess bond returns on the predicted excess returns, 84% of the observations are in the upper-right quadrant or in the lower-left quadrant. Moreover, the correlation between the *out-of-sample* predictions and the subsequent twelve-month excess returns is 0.57, much larger than the 0.26 correlation between these predictions and the subsequent monthly excess returns.

Exhibit 6 displays a time series plot of the monthly predicted excess returns and the subsequent twelve-month average excess returns. We also plot the time series of each predictor variable in Exhibit 7, to better identify the sources of fluctuations in expected excess bond returns.

EXHIBIT 6 ■ Predicted Excess Bond Return and Subsequent Realized Twelve-Month Excess Bond Return, 1965-1995

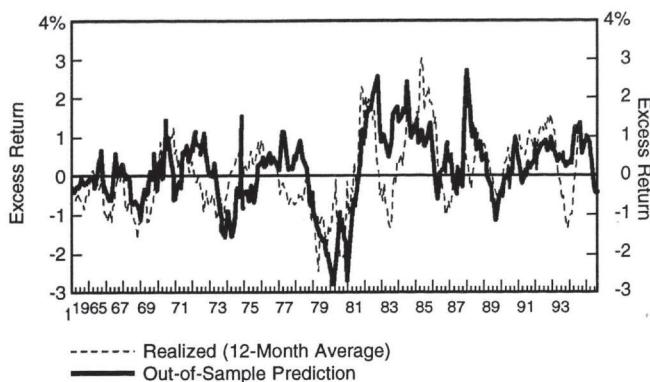
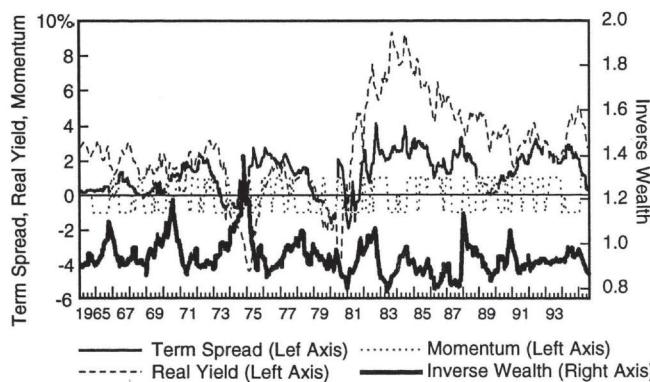


Exhibit 6 shows that the predictions track the movements in the realized bond returns reasonably well. Both series in Exhibit 6 are low in the 1960s and 1970s and exceptionally high in the 1982-1985 period, reflecting slow-moving changes in the real yield and in the term spread. Aside from these broad movements, both series exhibit apparent business cycle patterns: The predicted returns tend to increase during cyclical contractions such as those in 1970, 1974, and 1982.

Exhibit 7 shows that these increases in expected returns, as well as those in the aftermath of the 1987 stock market crash (when a recession was widely expected) and in the most recent recession (the 1990 Gulf War), coincide with inverse wealth spikes, that is, with poor stock market performance. These patterns

EXHIBIT 7 ■ Historical Levels of the Predictor Variables, 1965-1995



are consistent with our hypothesis that wealth-dependent risk aversion causes the required bond risk premium to vary over time with economic conditions. The negative excess return forecasts in the 1960s, 1973-1974, 1979-1980, and 1989 are difficult to interpret, however; most bond market participants think that the bond risk premium is always positive.

Finally, it is worth noting that the forecasting model is quite bearish at the time of this writing in August 1995. The excess bond return predictions have been negative since the end of May 1995, mainly because of the strong stock market (low inverse wealth) and the relatively flat yield curve (low term spread).⁹

III. INVESTMENT IMPLICATIONS OF BOND RETURN PREDICTABILITY

Exploiting Return Predictability by Using Dynamic Investment Strategies

Even if the predictor variables appear to have some ability to forecast excess bond returns in an out-of-sample setting, should investors care about these findings? For portfolio managers, the key question is whether investment strategies that exploit the return predictability produce economically significant profits.

To answer this question, we first describe the implementation of such dynamic investment strategies and then present extensive analysis of their historical performance. In particular, we compare their historical returns to the returns of static strategies that have a constant portfolio composition regardless of economic conditions.

One goal is to show that the information in the yield curve and in other predictors could have been used to enhance long-run returns. Another goal is to provide a tool kit to evaluate the future profitability of any forecasting strategy and to pose a set of critical questions (economic reason for success, stability of success, sensitivity to transaction costs and to risk adjustment) that investors should ask when faced with back-test evidence of an apparently attractive investment strategy.

The static strategies are called "always-bond" and "bond-cash combination." The first strategy involves always holding a twenty-year Treasury bond; the other involves always holding 50% of the portfolio in cash (one-month Treasury bill) and 50% in the twenty-year Treasury bond with monthly rebalancing.

The dynamic strategies adjust the allocation of the portfolio between cash and the twenty-year Treasury bond each month, based on the predicted value of next month's excess bond return. The two dynamic strategies are called "scaled" and "1/0."

The 1/0 strategy is simpler: It involves holding one unit of the twenty-year bond when its predicted excess return is positive and none when it is negative (thus, holding cash). This approach ignores information about the magnitude of the predicted excess return. The scaled strategy involves buying more long-term bonds, the larger the predicted excess return is. Specifically, investors should buy or short-sell the long-term bond in proportion to the size of the predicted excess return.¹⁰

Strategy returns are expressed in terms of an excess of the one-month bill return. For investors who have investable funds, each strategy's total return would be approximately equal to its excess return plus the one-month bill's return (which is the same number for all the strategies). For arbitrage traders who hold only

self-financed positions, the reported excess return can be interpreted as the total return of their "zero-net investment" position — to the extent that they can finance their position using the one-month bill rate.

Exhibit 8 shows, for each strategy, the annualized average excess return, the volatility of excess returns, and the Sharpe ratio. Note that a cash portfolio earns zero excess return by definition; it is equivalent to holding cash financed with cash. Therefore, the fifty-fifty bond-cash combination has exactly half of the excess return and volatility of the always-bond strategy.

The static strategies yield only insignificant excess returns over the sample period. In other words, long-term bonds and short-term bonds earn quite similar average returns.

The dynamic strategies perform much better. The scaled strategy earns almost 9% average annual excess return, while the 1/0 strategy earns about half that. Also the rewards to volatility (Sharpe ratios) of the dynamic strategies are much larger than those of the static strategies.¹¹

We can compare the performance of the dynamic strategies in Exhibit 8 with the performance of a dynamic strategy that uses only the information in the term spread. The scaled strategy would have earned 3.87% per year and the 1/0 strategy 2.96% per year if the out-of-sample forecasts had been based on the term spread alone. Comparison with the average returns in Exhibit 8 (8.64% and 4.16%) indicates that the marginal value of the other predictors has been substantial.

It is also worth noting that the scaled strategy would have earned 11.15% per year and the 1/0 strategy 4.94% per year if the predictions had been based on the in-sample estimates from the regression of excess bond returns on the four predictors. The difference between the performance of the in-sample and the out-of-sample forecasts may reflect the data-snooping bias.

Stability of the Predictive Relations

The analysis above shows, first, that over the past thirty years, our predictors have been able to forecast near-term bond returns and, second, that strategies that exploit such predictability have earned economically meaningful profits. Are these findings stable over time?

If the predictive ability and exceptional performance arise from a couple of extreme observations, we would be skeptical about the reliability of these findings. If observed relations are consistent across subperiods, we would think they are less likely to be spurious

EXHIBIT 8 ■ Performance of Self-Financed Dynamic and Static Investment Strategies, 1965-1995

Dynamic Strategies

Scaled Strategy

Average Excess Return	8.64%
Volatility	12.80
Sharpe Ratio	0.68

1/0 Strategy

Average Excess Return	4.16%
Volatility	7.92
Sharpe Ratio	0.53

Static Strategies

Always-Bond

Average Excess Return	0.67%
Volatility	10.42
Sharpe Ratio	0.06

Bond-Cash Combination

Average Excess Return	0.33%
Volatility	5.21
Sharpe Ratio	0.06

Note: Average excess return is the annualized average excess return of each strategy over the one-month bill, expressed in percent. Volatility is the annualized standard deviation of the excess return series. The Sharpe ratio is the ratio of the (annualized) average excess return to volatility.

and become more comfortable with expecting that the historical experience (good predictive ability and the dynamic strategies' exceptional performance) will be repeated in the future.

We look at three types of evidence: rolling correlations; a subperiod analysis of average returns; and the cumulative performance of various investment strategies.

Exhibit 9 shows that estimated rolling sixty-month correlations between the predictors and subsequent bond returns are not constant, but they are positive in most subperiods. In the 1990s, the real yield and momentum have had little forecasting ability, but both the term spread and inverse wealth have predictive correlations near 20%.

The combined predictor tends to have better forecasting ability than any of the individual predictors. Similar subperiod analysis shows that the frequency of correct prediction of the sign (+/-) of the next month's excess bond return is reasonably stable and near 60%.

Exhibit 10 reports the statistics from Exhibit 8 for three decade-long subsamples and for the 1990s subperiod. It is encouraging to see that the observed patterns are stable across decade-long subsamples. In particular, both dynamic strategies outperform the bond-cash combination strategy by at least 200 basis points in all subperiods.

The most informative way to display the stability of a predictive relation is to plot the cumulative wealth of an investment strategy that exploits the pre-

EXHIBIT 9 ■ Rolling 60-Month Correlation of Various Predictors with Subsequent Excess Bond Return, 1965-1995

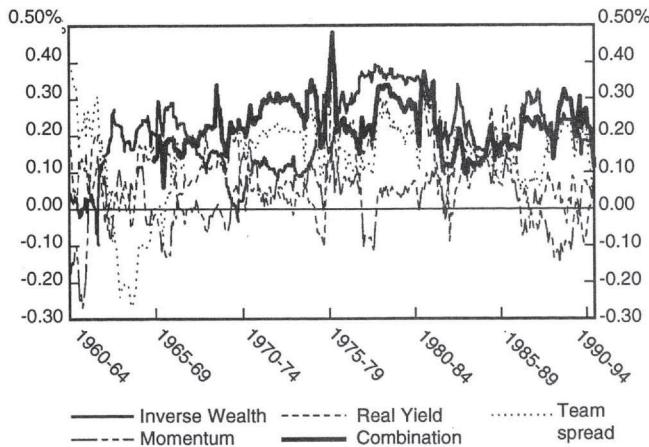


EXHIBIT 10 ■ Subperiod Performance of Various Investment Strategies

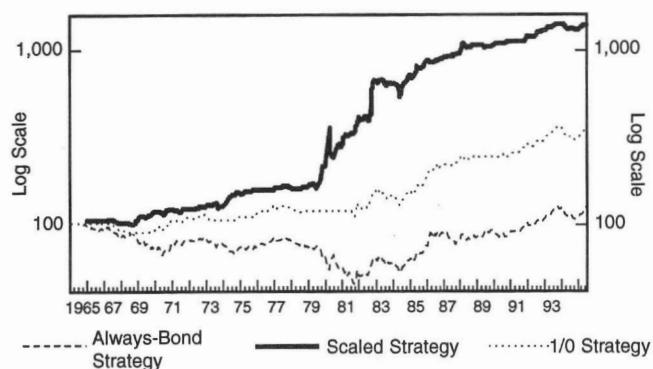
	1965-1974	1975-1984	1985-1994	1990-1995
Dynamic Strategies				
Scaled Strategy				
Average Excess				
Return	4.09%	15.29%	6.40%	5.30%
Volatility	5.65	20.30	7.32	4.80
Sharpe Ratio	0.72	0.75	0.87	1.10
1/0 Strategy				
Average Excess				
Return	0.89%	3.20%	7.27%	6.73%
Volatility	5.56	8.92	8.66	7.79
Sharpe Ratio	0.16	0.36	0.84	0.86
Static Strategies				
Always-Bond				
Average Excess				
Return	-3.13%	-1.68%	5.62%	5.25%
Volatility	8.35	12.36	10.05	8.28
Sharpe Ratio	-0.38	-0.14	0.56	0.63
Bond-Cash Combination				
Average Excess				
Return	-1.57%	-0.84%	2.81%	2.72%
Volatility	4.18	6.18	5.03	4.14
Sharpe Ratio	-0.38	-0.14	0.56	0.63

dictive relation. Such a graph shows how the profits from the strategy grow over time.

Note that the cumulative wealth of an ideal perfect foresight strategy would never decline; moreover, it should also rise faster than the cumulative wealth of any competing strategies. Alternatively, we can plot the relative performance of two investment strategies, and again, the line representing a perfect foresight strategy should always be rising (or flat if it matches the performance of the other strategies).

Exhibit 11 shows the cumulative wealth growth of both dynamic strategies and the always-bond strategy (plotted on a log scale where constant percentage growth produces a straight line). Because the lines cumulate each strategy's monthly returns in excess of cash (the one-month bill), we also can interpret these lines as relative performance versus cash. Exhibit 12 measures the relative performance of the two dynamic strategies versus a more realistic benchmark, a fifty-fifty combination of cash and the long-term bond.

EXHIBIT 11 ■ Cumulative Wealth Growth from Three Self-Financed Strategies, 1965-1995



These graphs show that the dynamic strategies have had a consistent ability to outperform the static strategies. The scaled strategy earns very high returns in the late 1970s and early 1980s by short-selling the long-term bond during the bear market. During the subsequent bull market, the dynamic strategies have earned returns similar to as the static bondholding strategy. This result must be viewed as satisfactory, because this bull market has been exceptionally strong and long, making long-term bond returns a difficult target to beat.

Exhibits 11 and 12 show that the dynamic strategies never underperform the benchmark static strategies for an extended period. And what about the

recent experience? The dynamic strategies have outperformed the static strategies in the 1990-1995 period — see Exhibit 10 — but in 1994 both dynamic strategies underperformed the cash-bond combination because they remained in long-term bonds throughout a period of rising rates.

Other Critical Considerations

The backtest results suggest that bond investors could enhance their performance substantially by exploiting the forecasting ability of the term spread, the real yield, inverse wealth, and momentum. Historical success, however, does not guarantee future success. Any reported findings of apparently profitable investment strategies should be subjected to a set of critical questions.

We have mentioned the important concern about a data-snooping bias — we use out-of-sample forecasts, we restrict the predictors to economically well-motivated variables, and we ensure that the observed findings are relatively stable across subperiods. Other reservations include the sensitivity of the findings to transaction costs and to risk adjustment.

Transaction costs will reduce the profitability of any investment strategy, although government bonds have such small transaction costs for institutional investors that their impact on the reported returns should be small. In particular, the results of the 1/0 strategy are hardly affected because this strategy involves very infrequent trading — on average 1.5 trades per year. The scaled strategy is more transaction-intensive, and it also involves short-selling. Thus, the reported results are somewhat exaggerated.

The dynamic strategies offer higher returns than the static strategies, but they excel even more when the comparison is between risk-adjusted returns. First, if risk is measured by the volatility of returns, the Sharpe ratios in Exhibit 8 provide a risk-adjusted comparison. The volatility of the scaled strategy is higher than that of the static bond strategy, but its reward-to-volatility ratio is more than ten times higher. The volatility of the 1/0 strategy is lower and the average return higher than that of the static always-bond strategy.

Second, if investors are concerned with downside risk, the dynamic strategies look even better. The historical success of these strategies partly reflects their ability to avoid long-term bonds during bear markets. For example, we can infer from Exhibit 5 that the 1/0 strategy underperforms cash in only 26% of the months

EXHIBIT 12 ■ Dynamic Strategies' Relative Performance versus Bond-Cash Combination, 1965-1995

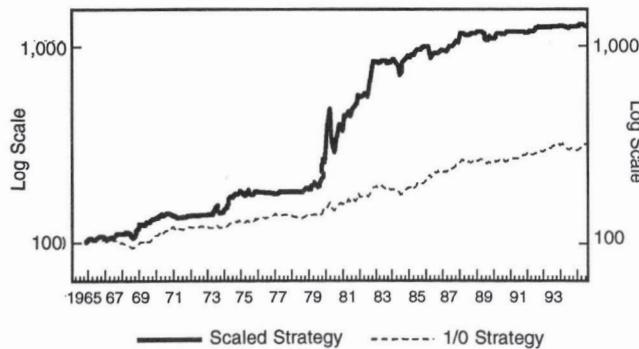


EXHIBIT 13 ■ Impact of the Forecast Signal's Strength on Return Predictability, 1965-1995

	$f < -1$	$-1 < f < -0.5$	$-0.5 < f < 0$	$0 < f < 0.5$	$0.5 < f < 1$	$f > 1$
Frequency of Correct-Sign Predictions	65%	63%	65%	53%	57%	67%
Average Excess Return	-15.68	-3.80	-8.57	3.92	4.15	15.66
Number of Months of Total (%)	9	11	19	25	21	15

Note: f is the beginning-of-month out-of-sample forecast of the twenty-year bond's monthly excess return, expressed in percent per month. Average excess return is the annualized average of the twenty-year bond's monthly excess returns within each subsample, in percent.

in the sample (outperforming it 35% of the time and matching its performance 39% of the time when the predicted excess return is negative and the strategy involves holding cash).

If the return predictability reflects a time-varying risk premium, however, it is possible that the abnormally high returns of the dynamic strategies reflect only a fair compensation for taking on additional risk at times when either the amount of risk or risk aversion is abnormally high.

In spite of the apparent attractiveness of the dynamic strategies, few investors have tried to systematically exploit the predictability of bond returns. For investors who venture to do that, this fact is good news. The profit opportunities are not likely to be "arbitraged away" any time soon.

One major reason is that these strategies are not riskless arbitrages — they involve a lot of short-term risk since the forecasts are wrong 40% of the time. Nonetheless, sixty-forty odds are attractive in competitive financial markets.

What could make investors forgo the exceptionally favorable odds that the dynamic strategies offer? Here are some possible explanations:

- Many investors prefer the more subjective interest rate forecasting approach even if its track record is rarely good. Other investors believe that market fluctuations cannot be predicted; thus, they do not want to take any market-directional positions. Such investors would attribute our predictability findings to data snooping or to events that the market was expecting but that were not realized during the sample period.

- The potential losses from such strategies may loom larger than the potential gains. The short-term risk in the dynamic strategies may expose portfolio managers to substantial career risk even if the strategies are likely to outperform in the long run.¹² Moreover, the losses may have a tendency to occur at especially unpleasant times.¹³ The dynamic strategies' high expected return may be a reward for such discomforts.

Even if investors find a dynamic strategy too mechanical or too risky to be used systematically, they may want to use it *selectively*. For example, they may want to use the strategy only when the signal is very strong.

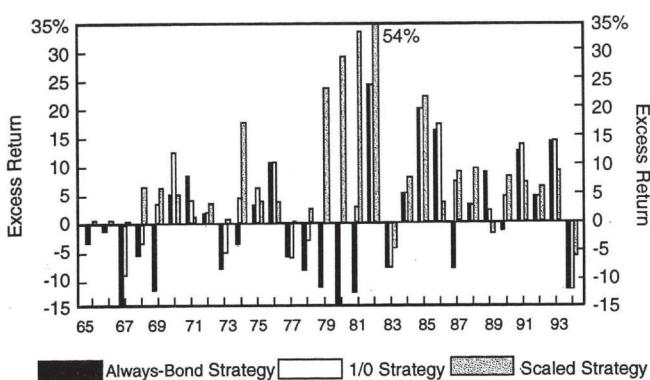
What do historical data say about such an approach? The fact that the scaled strategy outperforms the 1/0 strategy suggests that the magnitude of the predicted excess return provides valuable information beyond the sign.

Exhibit 13 shows evidence on whether the return predictions become more reliable when the forecast deviates much from zero. It also reports the average returns at different levels of predicted excess returns. We can see that the frequency of correct sign forecasts is only weakly related to the absolute value of the forecast. The average excess returns show clearer patterns: large negative values when the predictions are very negative, and large positive values when the predictions are very positive.

Impact of Investment Horizon

We have suggested that the dynamic strategies may involve an unacceptably high risk of short-term underperformance. The long-run performance of these

EXHIBIT 14 ■ Annual Excess Returns of Three Self-Financed Strategies, 1965-1994



strategies should nevertheless make them very appealing for investors who can afford to take a long investment horizon.

How long a horizon is long enough for investors to be confident that these strategies outperform cash and/or bonds? Recall that the out-of-sample forecasts of next month's excess bond return have a correct sign in 61% of the months in the sample (see Exhibit 15). Increasing the investment horizon from one month makes the dynamic strategies look better and better.

For example, Exhibit 14 shows that the scaled strategy outperforms the always-bond strategy in twenty-three calendar years out of thirty in this sample, and the 1/0 strategy underperforms the always-bond strategy in only two calendar years (and matches it in ten other years when it kept holding the long-term bond).

Exhibit 15 shows — for various horizons — the frequency with which the dynamic strategies outper-

form or match cash, the long-term bond, or both. The longer horizon numbers are based on overlapping monthly data.¹⁴

Focusing on the toughest comparison, the dynamic strategies outperform *both* bonds and cash in roughly 60% of the one-year periods in the sample. For the three-year horizon, the frequency increases to about 80% of the sample, and for the five-year horizon to 92%.

In a way, the dynamic strategies would have provided a free outperformance option for long-term investors. Exhibit 16 illustrates this point by showing the rolling thirty-six-month excess return for each strategy. If the dynamic strategies outperform both cash and bonds, their excess returns should lie above the excess returns of cash (the zero line) and the always-bond strategy. This is roughly what we see in the graph.

We conclude that although historical analysis provides no guarantee about the future, and although backtest results are rarely achieved in real-world investments, the odds in favor of the dynamic strategies appear excellent for three- to five-year horizons.

IV. CONCLUDING REMARKS

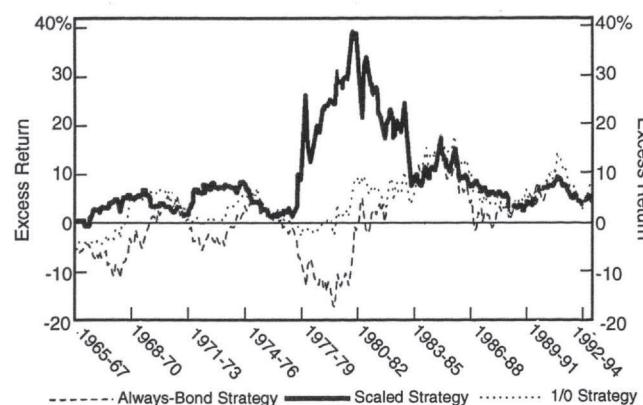
We have shown that long-term bond returns are predictable. A set of four predictors — yield curve steepness, real bond yield, recent stock market performance, and bond market momentum — are able to forecast 10% of the monthly variation in long-term bonds' excess returns.

When making inferences about yield curve behavior, bond market analysts should not assume that the bond risk premium is constant over time. The bond risk premium may be small, on average, but we can identify in advance periods when it is abnormally large

EXHIBIT 15 ■ Impact of the Horizon Length on the Strategy's Success Rate, 1965-1995

Horizon	Scaled Strategy			1/0 Strategy		
	Beat/ Match	Beat/ Bond	Beat/ Both	Beat/ Match	Beat/ Bond	Beat/ Both
	Cash	Bond	Both	Cash	Bond	Both
Month	61%	56%	35%	74%	86%	61%
Quarter	68	61	46	71	85	58
Year	84	69	58	73	87	64
Three Years	99	86	85	82	91	73
Five Years	100	92	92	94	97	92

EXHIBIT 16 ■ Rolling 36-Month Excess Returns of Three Self-Financed Strategies, 1968-1995



or small. A forecasting model gives us an estimate of the near-term bond risk premium—but even the best models' estimates are subject to various errors.

Nevertheless, such models can be valuable tools for long-term investors. We find that dynamic investment strategies that exploit bond return predictability have consistently outperformed static investment strategies over long investment horizons.

There are many ways to implement investment strategies that exploit return predictability. We describe two dynamic strategies (scaled and 1/0) that shift funds between cash and a long-term bond on the basis of the sign and the magnitude of the return prediction.

An alternative way to implement the strategy is through active duration management using on-the-run

EXHIBIT 17 ■ Implementing the Strategy: From Return Predictions to Target Portfolio Durations

Predicted Excess Bond Return	Target Duration for an Aggressive Investor	Target Duration for a Moderate Investor	Target Duration for a Conservative Investor
Negative	-2	0	3.0
Low	1	2	3.5
Average	4	4	4.0
High	7	6	4.5
Very High	10	8	5.0

Treasury bonds or bond futures. An investor could modify portfolio duration dynamically on the basis of the magnitude of the return prediction. The range of durations would depend on the investor's risk aversion level and confidence in the proposed strategy.

Exhibit 17 gives an example of how various investor types (aggressive, moderate, conservative) with a neutral benchmark duration of four years could vary their portfolio's target duration with economic conditions. Of course, more conservative implementation would reduce the potential for return enhancement.

The analysis in this article focuses on the predictability of the excess return of a twenty-year U.S. Treasury bond using four predictor variables. Obviously, the analysis could be extended in various directions:

- One might improve the forecasts by using a broader set of predictors, or by combining their information in a more sophisticated way than a simple linear regression. Even our small set of predictors, however, may have more robust forecasting ability in an out-of-sample setting than a broad predictor set would. We have not found other predictors that consistently and significantly improve the forecasting ability of our four-predictor model.
- One can examine the return predictability over a shorter investment horizon than one month. The predictors we use may be too slow-moving for short-term traders. They often prefer to trade on their fundamental views, or on momentum and overreaction effects (price trends and reversals), or on other technical factors (supply effects and portfolio flows). It might be a good idea to subject even these trading approaches to the performance evaluation proposed here.
- One can examine the return predictability of other government bonds. We show elsewhere that our predictors can also forecast the excess returns of shorter-maturity bills and bonds in the United States, and that similar variables forecast international bond returns—see Ilmanen [1994, 1995, 1996c].
- One can examine the predictability of the relative performance of various bond market sectors and markets (changes in the yield spreads across maturities, across market sectors, and across countries). In this article, we combine the information in the term spread and other predictors to improve our forecasts of excess bond returns. In a similar way, we could

combine the information in mean-reverting yield spreads and in other predictors to develop better relative value indicators. The tools presented here can also be used to evaluate the forecasting performance of various relative value indicators.

ENDNOTES

This article was originally published as a Salomon Brothers research report titled "Forecasting U.S. Bond Returns: Understanding the Yield Curve: Part 4." The author thanks Lawrence Bader, Eduardo Canabarro, Deepak Gulrajani, Robin Grieves, Ray Iwanowski, and John Liew for their helpful comments.

¹We define the *bond risk premium* as the near-term (say, one-month) expected return of a long-term government bond in excess of the return of the near-term riskless asset (say, the one-month Treasury bill). The *long-run* bond risk premium is the long-run average expected return of a long-duration strategy in excess of the long-run average expected return of a short-duration strategy, and it may differ from the (near-term) bond risk premium if the latter is abnormally high or low. Our definition of bond risk premium encompasses any expected return differential between a long-term bond and the near-term riskless bond, whether it is actually related to risk or to some technical factor. For this reason, we often call the bond risk premium, more neutrally, the "expected excess bond return."

²Moving average trading rules are perhaps the most popular trend-following strategies among traders. Such strategies are profitable if the market moves more in trends than sideways within a trading range. Research in the 1960s and 1970s found that common technical trading rules do not consistently outperform buy-and-hold strategies, but more recent studies have shown that some trend-following strategies are profitable, especially in the foreign exchange market. See, for example, Silber [1994]. Momentum indicators are often viewed as indicators of the market sentiment. An alternative interpretation, which is consistent with economic theories and rational behavior, is that the trends in bond markets reflect slow declines (increases) in the bond risk premium that coincide with bull (bear) markets.

³More sophisticated ways have been developed to combine the information in several predictive variables. See, for example, Sorensen et al. [1991] and Mezrich [1994].

⁴We forecast the *excess* return rather than return for three reasons: 1) the former is a proxy for the realized risk premium of a long-term bond (because any asset's return can be viewed as the sum of the riskless return and a realized risk premium); 2) it corresponds to a return on a self-financed position; and 3) it is harder to predict (because we subtract the riskless return that is known at the time of forecasting).

This choice hardly affects the predictability findings because the correlation between returns and excess returns is 0.997. (We also could forecast the long-term rate changes, whose correlation with excess bond returns is -0.97.) Finally, we examine a twenty-year bond because it has a long historical return series; the main findings are similar if we examine a shorter history of a ten-year or a thirty-year bond. This similarity is not surprising because the returns of all long-term government bonds are highly correlated.

⁵A correlation coefficient measures how closely two series move together. Its possible values range from -1 to 1, where 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates the lack of any correlation. The square of the correlation coefficient, so-called R^2 , measures what part of the variability in a regression's dependent variable (say, excess bond return) is explained by the variation in the independent variables.

⁶The correlations may appear low to readers who are used to examining contemporaneous relations between bond returns and other variables. It is easy to *explain* a large part of the fluctuations in monthly excess bond returns, but it is very difficult to *forecast* those returns. Because most of the realized monthly excess returns are unexpected, even the optimal predictor will have a low correlation with the realized return. In fact, if the bond risk premium is constant over time, we should expect to see zero correlations between the excess bond return and its predictors, which are known at the beginning of the forecasting month.

⁷Predictability of U.S. bond returns is discussed in Keim and Stambaugh [1986], Fama and Bliss [1987], Fama and French [1989], Jones [1992], Campbell [1995], and Tyson [1995]. Predictability of international bond returns is discussed in Mankiw [1986], Bisignano [1987], and Ilmanen [1994, 1995, 1996].

⁸This approach still leaves one element of hindsight in the analysis: The predictors may have been chosen on the basis of their historical fit. It may be unrealistic to assume that bond analysts chose to focus on this particular set of predictors (out of many alternative sets) a long time ago. We have three answers to such criticism: 1) We can motivate our predictors with economic reasoning; 2) the relation between the predictors and future bond returns is reasonably stable. Data analysis could have alerted investors of these relations a long time ago; and 3) the ultimate test is the predictors' performance with "new" data.

⁹We reemphasize that the model produces only estimates, and that these estimates capture, at best, 10% of the future variation in excess bond returns. Thus, if *unexpected* economic news in the coming months lowers the market's rate expectations, the bond market may perform well in spite of the low risk premium. The model just signals that the expected return cushion in favor of long-term bonds is abnormally low, making these bonds vulnerable to bad news.

and to a rising risk premium.

¹⁰An example illustrates how the scaled strategy works. If the predicted bond risk premium (BRP) over the next month is 0, the scaled strategy involves buying no long-term bonds, just cash. If the BRP is 1%, the strategy involves buying one unit of the long-term bond, no cash. If the BRP is 2%, the strategy involves buying two units of the long-term bond by using leverage (borrowing cash). If the BRP is -1%, the strategy involves short-selling one unit of the long-term bond and investing the sale proceeds in cash. Because the scaled strategy often involves either leveraging or short-selling, it is much riskier than the 1/0 strategy.

¹¹Note that the scaling intensity used in the scaled strategy is arbitrary. More aggressive scaling factors would lead to higher average returns and higher volatilities. Fortunately, the Sharpe ratios do not depend on the scaling factor. Exhibit 8 shows that the scaled strategy has the highest Sharpe ratios.

¹²First, the dynamic strategy can make the portfolio differ significantly from most peer portfolios. Given frequent performance evaluations, short-term underperformance relative to a peer group often implies serious career risk. Therefore, portfolio managers may avoid the dynamic strategies because "it is better to lose conventionally than to gain unconventionally." Second, the dynamic strategy works quite slowly. Many traders prefer making many trades with immediate outcomes to making one trade with a delayed outcome. The former approach allows better diversification (across a large number of trades within a given period) — and smaller career risk — than the latter approach.

¹³For example, the dynamic strategies might systematically underperform during recessions when many investors find it particularly difficult to tolerate losses. Our empirical analysis, though, shows that the dynamic strategies tend to perform particularly well in "bad times." During cyclical recessions, as defined by the National Bureau of Economic Research, the always-bond strategy earns an annualized average excess return of 6.55%, while the scaled strategy earns 19.75% and the 1/0 strategy 9.08%. Similarly, both dynamic strategies tend to outperform the static strategies near business cycle troughs, in months when excess bond returns are negative, and following years of exceptionally poor bond market performance.

¹⁴For example, the evaluation at the five-year horizon compares the holding-period returns of various investment strategies between January 1965 and December 1969, February 1965 and January 1970, March 1965 and February 1970, and so on, until August 1990 and July 1995. There are a total of 307 overlapping five-year periods. (Recall that each strategy may involve monthly rebalancing. The dynamic strategies use the predictions to adjust the portfolio; the bond-cash combination is rebalanced to fifty-fifty shares; and

even the always-bond strategy requires occasional rebalancing so that the portfolio maturity does not deviate too much from twenty years.) The last row in Exhibit 15 reports how frequently — that is, in what part of the 307 five-year periods — the dynamic strategies beat or match the performance of cash, bond, or both.

REFERENCES

- Bisignano, J. "A Study of Efficiency and Volatility in Government Securities Market." *Bank for International Settlement*, 1987.
- Campbell, J.Y. "Some Lessons from the Yield Curve." *National Bureau of Economic Research* working paper #5031, 1995.
- Campbell, J.Y., and J.H. Cochrane. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." NBER working paper #4995, 1995.
- Chen, N. "Financial Market Opportunities and the Macroeconomy." *Journal of Finance*, 1991.
- Fama, E.F., and R.R. Bliss. "The Information in Long-Maturity Forward Rates." *American Economic Review*, 1987.
- Fama, E.F., and K.R. French. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics*, 1989.
- Ilmanen, A. "Does Duration Extension Enhance Long-Term Expected Returns?" *Journal of Fixed Income*, September 1996a.
- . "Market Rate Expectations and Forward Rates." *Journal of Fixed Income*, September 1996b.
- . "Time-Varying Expected Returns in International Bond Markets." Doctoral dissertation, Graduate School of Business, University of Chicago, 1994.
- . "Time-Varying Expected Returns in International Bond Markets." *Journal of Finance*, 1995.
- . "When Do Bond Markets Reward Investors for Interest Rate Risk?" *Journal of Portfolio Management*, 1996c.
- Jones, I.E. "Can a Simplified Approach to Bond Portfolio Management Increase Return and Reduce Risk?" *Journal of Portfolio Management*, 1992.

- Keim, D., and R. Stambaugh. "Predicting Returns in the Bond and Stock Markets." *Journal of Financial Economics*, 1986.
- Mankiw, N.G. "The Term Structure of Interest Rates Revisited." *Brookings Papers on Economic Activity*, 1986.
- Marcus, A.J. "An Equilibrium Theory of Excess Volatility and Mean Reversion in Stock Market Prices." *National Bureau of Economic Research* working paper #3106, 1989.
- Mezrich, J.J. "When is a Tree a Hedge?" *Financial Analysts Journal*, November–December 1994.
- Sharpe, W.F. "Investor Wealth Measures and Expected Return," in seminar proceeding Quantifying the Market Risk Premium Phenomenon for Investment Decision Making. The Institute of Chartered Financial Analysis, 1990.
- Silber, W.L. "Technical Trading: When it Works and When it Doesn't." *Journal of Derivatives*, Spring 1994.
- Sorensen, E.H., J.J. Mezrich, M.A. Beck, and G.G. Shkedy. "Anatomy of the Global Allocation Model." Salomon Brothers, Inc. September 1991.
- Tyson, D. "Estimation of the Term Premium," in T.S.Y. Ho, ed., *Fixed Income Investment: Recent Research*. Bunker Ridge, IL: Irwin, 1995.

The views and opinions expressed herein are those of the authors and do not necessarily reflect the views of AQR Capital Management, LLC, its affiliates or its employees. This document does not represent valuation judgements, investment advice, or research with respect to any financial instrument, issuer, security, or sector that may be described or referenced herein and does not represent a formal or official view of AQR.