Term Structure Information and Bond Strategies

María de la O González[a], Frank S. Skinner[b*] and Samuel Agyei-Ampomah[c]

[a] *Assistant Professor of Finance, Departamento de Análisis Económico y Finanzas, Universidad de Castilla-La Mancha, Facultad de CC Económicas y Empresariales de Albacete, Plaza de la Universidad, 1, 02071 – Albacete (Spain)*

[b] *Professor of Finance, Brunel University, Department of Economics and Finance, School of Social Sciences, Uxbridge, Middlesex, United Kingdom*

[c] *Senior Lecturer of Finance respectively, University of Surrey, Surrey Business School, Guildford, Surrey, United Kingdom*

Key Words: Term structure theory, Bond strategies, MPPM, Performance measurement, Simulation.

JEL Classification: C61; E43; G11; G12

[*]Corresponding author, Tel: +44 (0) 189 526 7948, Fax: +44 (0)1895 269786.
E-mail addresses: MariaO.Gonzalez@uclm.es (María de la O González), Frank.skinner@brunel.ac.uk (Frank S. Skinner) S.Agyei-Ampomah@surrey.ac.uk (Samuel Agyei-Ampomah) Any errors are the responsibility of the authors.

Term Structure Information and Bond Strategies

Abstract

We examine term structure theories by using a novel approach. We form bond investment strategies based on different theories of the term structure in order to determine which strategy performs best. When using a manipulation-proof performance measure, we find that consistent with prior literature, an active strategy that is based on time varying term premiums can indeed form the basis of a successful bond strategy that outperforms an unbiased expectation inspired passive bond buy and hold strategy. This is true, however, for an earlier time period when the literature first made this claim. In a later time period, we find that the passive buy and hold strategy is significantly superior to all active strategies.

This result is confirmed by statistical tests and it suggests that once it became known that an active strategy based on time varying term premiums could outperform a passive buy and hold strategy, the markets adjusted and arbitraged away this opportunity. Overall, it appears that the unbiased expectation hypothesis is the most likely explanation of the behaviour of the term structure during more recent times. This is because economically and statistically significant superior performance cannot be achieved if one uses information from the forward curve or the term structure as a guide to adjusting bond portfolios in response to changes in the term premium.

1 Introduction

Recent discoveries suggest that actively managing bond portfolios by using information in the forward curve or the term structure of interest rates can lead to superior performance. Cochrane and Piazzesi (2005) and Kessler and Scherer (2009) find that the forward curve can predict future bond returns, while Estrella and Mishkin (1997, 1998) and Ang et al. (2006) find that the slope of the yield curve can forecast future rates of interest. Ilmanen (1995, 1997), Ilmanen and Sayood (2002) and Papageorgiou and Skinner (2002) all find that active strategies that are based on or placed in combination with time varying information in the term structure can form viable strategies for bond investors. Yet none of this work has been able to accurately address whether active strategies that are based on information in the term structure and the forward curve can indeed outperform a passive buy and hold strategy because this research did not have access to the more recently developed manipulation-proof performance measure (MPPM) of Ingersoll, Spiegel, Goetzmann and Welch (2007). Instead, the only performance measurement techniques that were available to these researchers were static in nature and thus unable to adjust for the inherent dynamic nature of strategies that are based on time varying term premiums incorporated in both the forward and term structure.

Given this gap in the research, the purpose of this paper is to determine whether dynamic strategies that are based on time varying term premiums can indeed outperform a benchmark buy and hold strategy. Central to our investigation is the issue of how to measure performance. Active strategies deliberately attempt to transform the distribution of returns by minimizing downside and enhancing upside potential, thereby creating positive skewness in their attempts to enhance returns. In the meantime, some strategies may succeed or fail dramatically, leading to fat tails in the distribution. If either or both the skewness and kurtosis are of concern to investors, any performance measure should account for these extra moments in the distribution of returns.[1] Moreover, by their very nature active strategies are dynamic, so static performance measures are unable to capture their essential nature. Ingersoll et al. (2007) observe that using static performance measures to evaluate dynamic strategies can be misleading because portfolio managers can manipulate their strategies, whether deliberately or not, in order to score well on a wide variety of static performance measures even though the manager has no private information. The MPPM overcomes the shortcomings of all previous static performance measures for several reasons. It is time separable and therefore not subject to dynamic manipulation; it is concave, which means that one cannot manipulate one's score through leverage; and it is consistent with equilibrium while at the same time recognizing superior performance that is based on the exploitation of genuine arbitrage opportunities.

Given these measurement issues, we measure the performance of a variety of bond strategies that are inspired by information that is supposed to be contained in the term structure and in the forward curve by means of five distinctly different performance measures. The first is the traditional mean variance Sharpe ratio; the second adjusts for utility functions that account for a preference for positive skewness as well as mean variance; the third adjusts for tail risk; the fourth adjusts for tail risk, skewness, and kurtosis, as well as mean variance; and the fifth adjusts for the dynamic nature of active strategies. This fifth measure is the MPPM, which prevents manipulation of performance scores by adjusting the return distribution through dynamic trading.

[1] Positive skewness implies a bias for positive returns and so is thought to be desirable by investors. In contrast, excess kurtosis, which implies fat tails and therefore a higher likelihood for extreme adverse outcomes, is thought to be an undesirable statistical attribute.

Some idea of the challenges that are posed for the investigation of the performance of bond investment strategies can be obtained by inspection of Figure I. Here we see that US interest rates have varied tremendously over recent decades. Even the more stable ten-year Treasury yield varied from 9.09% on May 5, 1990 to 3.13% on June 13, 2003, while the slope of the Treasury yield curve has turned negative for three periods after January 1990. Clearly, as an asset class, Treasury bond investment returns can vary tremendously, which translates into great risks as well as potentially great rewards. Care must thus be taken in order to accurately capture this dynamic interest rate environment. Therefore, we make strenuous efforts to calculate returns as accurately as possible by: reinvesting coupons on the day on which they are paid, purchasing bonds at their daily closing prices, reinvesting the proceeds of sales of bonds at the Libor rates prevailing on the day on which those bonds are sold; accruing interest according to the well-known Treasury and Libor market conventions; and reducing the proceeds of bond sales for the extra transactions costs that are required by active bond strategies.

<<Figure I about here>>

To enhance the robustness of as well as to measure the statistical significance of our results, we simulate our strategies 1,000 times via the bootstrap method (see Davison and Hinkley 1997). This experiment replicates the reported results for all strategies, and it provides the data that is needed to assess the statistical significance of differences in performance. Based on these bootstrap simulations, a strategy based on time varying term premiums in the forward curve provides statistically significantly superior performance when compared to the pure expectations buy and hold strategy for all static measures of performance. However, once we measure performance using the dynamic MPPM, we find that, both for overall performance and all sub periods, the bond buy and hold strategy is significantly superior, at the 1% level, to the forward curve strategy. Interestingly, the MPPM finds that a time varying premium strategy based on information in the term structure is significantly superior to the bond buy and hold strategy, but only for the first half of our sample period. For the second half of our sample period, the bond buy and hold strategy is significantly superior to all other strategies. This pattern suggests that the time varying risk premiums discovered earlier in the literature have since been arbitraged away.

2 Literature review

The theoretical justifications for each strategy are based on a fundamental theory of interest rates. The expectation hypothesis asserts that forward rates are related to investors' expectations concerning future rates of interest and so form unbiased predictions of future interest rates. This theory implies that, because term premiums are constant, there is no particularly good time to invest at a given maturity. An investor should buy and hold bonds with a maturity that is the same as his or her investment horizon. Most of the literature, such as Cochrane and Piazzesi (2005), Ilmanen (1996), and Fama and Bliss (1987), categorically reject the pure expectations theory of interest rates as it is evident that term premiums do in fact vary.

Despite this research, the expectations hypothesis refuses to die. Froot (1989) finds that while the expectations hypothesis is rejected at the short end of the yield curve, some support for the hypothesis is found at the long end. Longstaff (1990) finds that, for technical reasons at least, time varying term premiums can still be consistent with the expectations hypothesis. De Bondt and Bange (1992) suggest that time varying term premiums may be the result of skeptical under-reaction to inflation forecasts. Longstaff (2000) shows that the viability of the expectations hypothesis is purely an empirical issue because in an incomplete market the existence of the expectations hypothesis does not imply arbitrage opportunities. More recently, Galvani and Landon (2011) find that investors who are interested in short time horizons are better off investing in short term bonds rather than attempting to capture term premiums by holding long term bonds for short periods. For these reasons, a simple buy and hold strategy is still viable in its own right and not just a "straw man" strategy to be used to benchmark the success of other bond trading strategies.

Meanwhile, there is considerable empirical support for a time varying term premium, which implies that investors should follow a more active strategy and shift their allocations in response to changing term premiums. Fama and Bliss (1987) and Hardouvelis (1988) find evidence that forward rates can predict future spot interest rates. Fama and Bliss (1987) attribute this to the mean reversion tendency of interest rates. Ilmanen (1997) finds that forward rates are upwardly biased forecasts of future rates of interest, thus implying that risk premiums are a more convincing explanation of the yield curve shape than are unbiased expectations. Estrella and Mishkin (1997) find that increases in the slope of the term structure are associated with increases in inflation, while Estrella and Mishkin (1998) and Ang et al. (2006) find that decreases in the slope of the term structure can indicate an increased likelihood of future recessions. Cochrane and Piazzesi (2005) find that the forward curve can be used to predict future Treasury bond returns. Kessler and Scherer (2009) extend this finding to six other bond markets. Interestingly, Kalev and Inder (2006) find unexploited information in the term structure that is not incorporated into expectations.

Given these findings, investors should adjust their bond holdings in response to changes in term premiums. For instance, according to Estrella and Mishkin (1997), when the term structure and therefore the term premium increases, inflation is expected to increase; this means that future interest rates are also expected to increase. This suggests that, in general, one should sell long-term bonds and buy short-term bonds. Ilmanen (1995, 1997) finds that such an active strategy can outperform a passive, unbiased expectation inspired buy and hold strategy.

Modigliani and Sutch (1966) suggest that investors are generally just as concerned with income as they are with capital risk so that interest rate decreases can be as damaging as interest rate increases for investors like pension plans and insurance companies who have definite investment horizons. For this reason, investors who have preferred habitats should immunize themselves by matching the Macaulay duration of their portfolios to their investment horizons, as this will optimally balance income and capital risk. Except for Van Horne (1980), however, the theoretical foundations of immunization have found little empirical support. Still, immunization remains a popular strategy in industry and variations of the immunization strategy have continued to attract academic interest in such recent papers as Soto (2001), Ventura and Pereira (2006), and Diaz et al. (2009).

3 Data and procedures

We intend to empirically investigate a simple buy and hold long term bond strategy and compare its performance to three different bond strategies (a term structure, a forward curve, and an immunization strategy) to see whether the statistically significant information that is contained in the term structure and forward curve can form the basis of a superior investment strategy. We chose to investigate the performance of nine-year bond portfolios, as nine years appears to be a likely candidate for a typical, long term, preferred habitat investment horizon. We note that Germany regularly issues ten year Treasury bonds (Bunds), and while most European nations choose a variety of maturities, there always seem to be issues of ten year maturity included in these nations' deficit-financing programs. Moreover, early in the twenty-first century when the US ran budget surpluses, the US Treasury stopped issuing bonds of different maturities but maintained an active ten-year Treasury note auction program. It therefore appears that investors have a taste for ten-year bonds from which it will always be possible to form a nine-year, immunized bond portfolio.[2]

In making this investigation, we collect the daily closing bid price as well as the issue date, maturity date, coupon rate, day count convention, and ISIN number of all Treasury bonds that are available in Bloomberg as of May 1, 2007, 452 in all.[3] Based on daily closing prices, day count convention, coupon rate, and maturity and issue dates, we calculate the yield to maturity, Macaulay duration, and accrued interest for each trading day. As our proxy for the rate of return that is available for short-term investment, we collect from Bloomberg the three-month Libor rates. Because we know the day count conventions, we are able to calculate the implied Libor price for each trading day.

To be sure that the bond strategies are examined in a unified setting, all bond strategies begun on the same date always use the same bonds when invested in the

[2] We explored the possibility of investigating other investment time horizons, but over the seventeen years covered by this paper we found gaps in the maturities available. For example, we could not find any bonds of between five and six year's maturity between January 1990 and July 1993. Therefore, it was difficult to start a five year immunized portfolio during these 31 months without forming the portfolio from broadly divergent maturities, thereby incurring a significant risk of immunization failure. This would have heavily biased our results against the immunization strategy.

[3] To avoid survivorship bias, we include all the bonds that are available, including bonds that have matured prior to May 1, 2007.

bond market. To be eligible for selection, each bond must have nine years of continuous daily bid prices as of the date we form the initial portfolio. Typically, we find that at least fifty bonds meet this criterion at any given time. To reduce idiosyncratic risk, we select a total of six bonds from the bonds that are eligible as of the portfolio formation date. At each formation date, four portfolios are formed, one corresponding to each of the strategies—bond buy and hold (bond BH), immunization, and time varying slope premium and forward curve—all of which use the same six bonds. That way, the asset selection decision of each of the four strategies is held constant so that we can isolate the effect of the timing decision.

The choice of which bonds to include in these portfolios is not arbitrary. One bond must have a maturity as close as possible to nine years, and a second bond must have a Macaulay duration as close as possible to nine years. Two other bonds must have a Macaulay duration greater than, but still as close as possible to, nine years, and the final two must have a Macaulay duration that is less than, but still as close as possible to, nine years. We follow this selection procedure so that it is always possible to form an immunized portfolio with little risk of immunization failure because our bond portfolios will always be composed of bonds with durations similar to the overall portfolio duration. Bierwag (1979) and Fong and Vasieck (1984) note that the risk of immunization failure increases for portfolios that are composed of bonds with durations radically different from their portfolio durations.

On the first working day of each month from January 2, 1990 to April 1, 1998, we select six bonds from the fifty or so eligible bonds that are available on each date. For the bond BH strategy, we invest $100 million in equal dollar amounts in each bond and hold those bonds until nine years later. During the nine years that we run each bond buy and hold strategy, we reinvest all coupons in equal amounts in the same six bonds. That is, if a coupon payment of $3 million is paid on a given day, then $500,000 is reinvested in each of the initial six bonds at bid prices, including accrued interest, that prevail at the end of that day.

At the end of each nine-year investment horizon, the bonds are sold at the prevailing daily closing price plus accrued interest. We then annualize the nine-year holding period return for each of the 100 nine year holding periods that are formed monthly between January 2, 1990 and April 1, 1998 and ending between January 2, 1999 and April 2, 2007. These returns are used to measure the mean, standard deviation, skewness, and kurtosis of the buy and hold strategy's return distribution.

The immunization strategy is identical to the buy and hold strategy except that we do not use equal weights for the reinvestment of coupons or for the formation of the initial portfolio. Specifically, bond weights are chosen to make sure that the initial bond portfolio has a Macaulay duration that is equal to nine years. Coupons are reinvested to maintain a duration of the bond portfolio that is equal to the remaining time horizon, whereas all other bond strategies reinvest the coupons in equal dollar amounts in the six bonds that comprise the portfolio. Note that the immunized portfolios are rebalanced each day a coupon payment is made rather than periodically. Because interest rates fluctuate widely throughout the sample period, daily rebalancing forms an important innovation that ensures that our holding period returns are measured as realistically and accurately as possible.

The portfolios for active strategies are formed from the same bonds as the bond BH and immunization strategies. Like the bond BH strategy, these portfolios are equally weighted. However, when using active strategies, if term premiums vary it will be vital to determine when the investor should switch from holding bonds to money market instruments and *vice versa*. Also, when returning to the bond market from money market instruments, the active strategies will return to the same six bonds that initially comprised the portfolio at prevailing market prices including accrued interest.

We implement two variations of the time varying term premium strategy. The first is the slope premium strategy (see Estrella and Mishkin 1997), which assumes that say an increase in the slope of the term structure contains useful information concerning the likelihood of higher future inflation and/or real economic activity and consequently that increases in the slope of the term structure are related to increases in longer-term interest rates. Therefore, if the slope of the term structure increases one should sell bonds and invest in money market instruments. The second time-varying premium strategy is the forward curve strategy (see Cochrane and Piazzesi 2005), which asserts that the forward curve predicts the expected return on long-term bonds. When the predicted excess return on long term bonds is positive, one should hold bonds; when the predicted return is negative, one should sell bonds and invest in money market instruments.

The slope premium strategy predicts whether long-term yields will rise or fall. This zero, one specification of the forecast naturally leads to the probit model.

$$P(Y_{t+h} = 1) = F(\text{Constant} + \text{Change in Slope}_t) \quad (1)$$

The above equation says that when making a forecast at date *t*, the probability that the long-term yield Y in h periods in the future will either rise (Y_{t+h} = 1) or fall (Y_{t+h} = 0) is a function F of a constant and the change in the slope of the yield curve. The above equation is estimated via a maximum likelihood probit regression rather than OLS since the dependent variable is dichotomous, only being able to take the value of one or zero. An important feature of how we implement (1) is that the probit forecasts are out of sample. This means that we only use information that is available in the market at date t in order to make forecasts that are h periods ahead.

We chose to test the ability of (1) to forecast the direction of future long term rates of interest using daily three month Treasury yield and ten year constant maturity yield data as published by the Federal Reserve Bank of New York (see Table H15). The slope is measured as the spread between the ten year Treasury rate, the direction of which is being forecast, and the three month Treasury yield. The dependent variable is equal to one if the long term yield rose for that day; it is equal to zero if the long term rates decreased or remained the same for that day. Along with a constant, the probit model uses the change in the slope as of date t in order to forecast the direction of the long term yield for the next day. As is standard practice in implementing probit regression forecasts, we interpret a probit probability forecast of greater than 0.5 as predicting that the long term rate will rise and a probit probability forecast of 0.5 or less as predicting that the long term rate will fall.

We use six months of daily data, from July 3, 1989 until December 29, 1989, to forecast whether ten year rates are expected to rise on January 2, 1990. We then roll the probit regression forward one more day by adding January 2, 1990 in order to forecast whether the ten year rate of interest will rise on January 3, 1990. We continue to roll the probit regression forward day by day until March 30, 2007.

The results suggest that the probit model does in fact contain some forecast ability according to the test statistic developed by Henriksson and Merton (1981). Let η_1 be the number of successful predictions of an increase in interest rates and η be the number of predictions both successful and unsuccessful of an increase in interest rates. Additionally, N_1 is the population number of increases in interest rates and N is the total population size, including all increases and no increases in the ten-year rate of interest. Then the expected number of correct predictions of interest rate increases that are merely due to chance is[4]

$$E(\eta_1) = \frac{\eta \times N_1}{N} \tag{2}$$

Since $N_1 \approx N/2$, Lehmann (1975) theorem 19 implies that the normal distribution will be a good approximation of the above statistic.[5] Therefore the variance of (2) is

$$\sigma^2 = \frac{\eta_1 N_1 (N - N_1)(N - \eta)}{N^2 (N - 1)}$$

Using these statistics, one forms the usual t-test as follows:

$$t = \frac{\eta_1 - E(\eta_1)}{\sigma}$$

Table 1 reports that the probit model obtains an excess number of successful out of sample predictions of an increase and no increase in the ten year rate of interest that is statistically significant at the 1% level. Clearly, the probit model has some degree of forecast ability but whether this model can form an economically significant trading strategy is an issue this paper will attempt to resolve.

---

[4] We also calculate the above statistic for the expected number of correct predictions of no increase in interest rates by replacing η_1 with $(1- \eta_1)$, the number of times interest rates are successfully predicted not to increase, and replacing η with $(1-\eta)$, the number of times interest rates did not increase.
[5] This condition is met because, of the 4,336 observations, 1,906 or 44% are increases in interest rates and 2439 or 56% are decreases in interest rates.

<<Table 1 about here>>

As the interest rate process is stochastic, a blind application of the probit decision rule will result in far too many trades—1,969 in all. Instead, one must seek assurance that interests rates are on a mean-reverting trend. Therefore, we decide to switch from bonds to Libor money market instruments when, during any ten trading day period, the probit model predicts a rise in long term interest rates on seven or more out of those ten trading days; we then switch back to bonds when the probit model predicts a fall in long term interest rates on seven or more out of the ten trading days.[6]

The passive bond BH strategy has the great advantage of not requiring any transaction costs for turnover; that is, there are no costs associated with selling the entire bond portfolio and reinvesting in the money market and *vice versa*. Active strategies such as the slope premium strategy require a complete portfolio turnover once the decision to move from bonds to the money market is made. Therefore we would bias the results in favour of active strategies if we did not adjust for the impact of these additional transaction costs. It is impossible to locate a precise measure of transaction costs all of the time because we sometimes have only bid prices. However, there are always a few bonds of close to ten years maturity that include both bid and ask prices. The average difference between the bid and ask prices of these bonds is \$0.058 per \$100 face value.[7] Using a bid-ask spread of \$0.06 per \$100 means that it will cost \$60,000 to sell or buy \$100 million at face value. Knowing the dates that the

[6] When experimenting with a range of cutoffs in order to decide to switch from bonds to Libor loans and *vice versa*, we find that the results are qualitatively similar.
[7] The bid-ask spread ranged from a low of \$0.011 to a high of \$0.1118, with a median of 0.0601 per \$100.

slope premium strategy requires for trading, we can therefore estimate the transaction costs of running this strategy and measure the net return. All of our results for the active strategies are reported using these net returns.[8]

For the forward curve strategy, we regress the one year excess holding period return of the ten year Treasury bond *x* months in the future on the forward swap curve via the following regression:

$$R_{t+x}^{t+x+1} = \gamma_0 + \gamma_1 S_{1,t} + \gamma_2 F_{1,t} + \gamma_3 F_{2,t} + \gamma_4 F_{3,t} + \gamma_5 F_{4,t} + \gamma_6 F_{6,t} + \gamma_7 F_{9,t} + \varepsilon_t \qquad (3)$$

The dependent variable R is the one year (t +x to t+x+1) return on the benchmark ten year Treasury bond x months from the current date t less the date t+x three-month t-bill interest rate. Unlike Cochrane and Piazzesi (2005), we experiment with a range of forecast horizons x from one month to twelve months. The independent variables are the current time t one year swap rate S_1 and the one year F_1, and the two year F_2, three year F_3, four year F_4, six year F_6, and nine year F_9 forward rates, each of one year's maturity. These are derived from the one, two, three, four, five, seven, and ten year swap rates at the current time t. We use swap rates rather than Treasury rates in order to estimate the forward rates for two reasons. First, swap rates are benchmarked off the Treasury yield curve, and second, active traders dominate the swap market. For these reasons, swap rates more accurately reflect current bond market information given that swap rates are more frequently updated for current market conditions than the less actively traded secondary market for Treasury bonds. The ten year benchmark and swap rates are collected from DataStream and the three month t-bill rates are taken from the Federal Reserve from

[8] Typically the net return is 2 to 3 basis points lower than the gross annualized holding period return.

Table H15. We estimate (3) for forecast windows x of one, three, six, nine, and twelve months, the results of which are reported in Table 2.

<<Table 2 about here>>

For a one month forecast horizon, we obtain an R^2 similar to Cochrane and Piazzesi (2005). However, the R^2 figures decline from a fairly high 45% for a one-month window to a low of 16% for windows of nine months or longer. Consequently, we only report the forward curve strategy for a one month forecast window, as longer horizons obtain much poorer and less interesting results.[9]

Similar to the previously explained slope premium forecasts, we make sample predictions by performing rolling regressions on (3) until March 30, 2007. Specifically, we start by running (3) from July 3, 1989 until November 30, 1989 in order to estimate the parameters of the regression. We use these parameters along with the December 1 values of the forward curve to forecast whether ten year bond returns are expected to be positive one month later on January 2, 1990. We then repeat this procedure each day until March 1, 2007. We continue to invest in bonds (money market instruments) when the predicted holding period return on ten year Treasury bonds in one month's time is positive (negative) for ten consecutive days.[10]

4 Methodology and performance measures

[9] For the sake of brevity, we choose not to report the results for longer forecast windows, but they are available from the corresponding author upon request.

[10] When we experiment with a variety of rules to decide to switch from bonds to Libor loans and *vice versa*, we find that the results are similar (see footnote 6).

We calculate returns by running nine year bond BH, slope premium, forward curve, and immunization strategies from January 2, 1990 to April 2, 2007. This involves calculating 100 returns for each of the four different strategies.[11]

Figure II reports the time pattern of excess returns calculated as the annualized nine year holding period return, including the reinvestment of coupons and deducting for the transactions costs of portfolio turnover, less the average three month Treasury bill yield. Evidently, nine year excess returns have been decreasing for all investment strategies. Also, there appear to be two distinct cycles embedded in this downward trend, the latter cycle commencing in March 2003. We later check on the robustness of our results by examining the performance of all strategies for these two sub-periods. Figure II shows that the excess returns of the various investment strategies track each other fairly closely when no one strategy clearly dominates.

<<Figure II about here>>

The next step is to examine the distribution of the returns. Table 3 measures the mean, standard deviation, skewness, and kurtosis of the returns for the four bond strategies and for two equally sized sub-periods. Our measures of the mean, standard deviation, skewness, and kurtosis of these returns are annualized measures that relate to the whole nine year time horizon and are based on 100 replications of each strategy. It is appropriate to assess the standard deviation and other moments of the distribution over the entire investment horizon when assessing the performance of

[11] The return algorithm typically takes 30 minutes to complete for each of the 100 calculations as first the algorithm must select bonds and purchase them at prevailing market prices, then monitor daily the active management decision rules and coupon payment dates. If a coupon is paid or a decision rule trigger is encountered, bonds or Libor loans are bought or sold at prevailing market prices that include calculated accrued interest. These amounts are added to the investment and then the process continues until nine years or approximately 2,250 trading days have elapsed.

long term bond strategies because the standard deviation of bond returns from actual bond portfolios systematically decreases as the underlying bonds mature.

<<Table 3 about here>>

We can see that, consistent with Figure II, the average mean returns do not appear to be very different amongst the various strategies (the range of returns amongst the four strategies is 38 basis points). However, this calculation tends to understate the economic significance of these differences because the largest actual difference of total dollar return for the same time period is more than $28 million. This difference occurs between the slope premium strategy ($245 million) and the immunization strategy ($217 million). Moreover, Table 4 finds that these differences in mean return are often statistically significant when compared to the bond BH strategy. For example, the forward curve strategy has a return that is 28 basis points lower than that of the passive bond BH strategy, and this difference is statistically significant at the 1% level using the standard errors of Newey and West (1987).

<<Table 4 about here>>

Table 3 further reports that the standard deviation is mostly consistent with the inherent risk that is posed by each strategy. The active slope premium strategy has a higher standard deviation and the defensive immunization strategy has a lower standard deviation than the bond BH strategy. The forward curve strategy is an exception. It is active, yet it has the lowest standard deviation of all of the strategies.

Looking at other moments of the empirical distribution, we observe that all strategies are positively skewed but that the forward curve strategy is platykurtic. We also find that the active forward curve strategy has more positive skewness and the active slope premium strategy less positive skewness than the passive bond BH strategy. This suggests that at least the active forward curve strategy was successful in making a favourable transformation of the distribution of returns. Finally, we note that these findings generally repeat across the sub-periods even though the first half sub-period is much more volatile.

While the above results are interesting, it is difficult to reach any general conclusion regarding the performances of these strategies when examining each of the four moments of the distribution one by one. What is needed is a performance statistic that accounts for additional moments within the distribution.

Zakamouline and Koekebakker (2009) derive adjusted Sharpe ratios by considering investors' preferences for higher moments of the distribution within the expected utility framework. Their generalized Sharpe ratio that adjusts for investors who have a preference for positive skewness is

$$\mathrm{ASSR} = \mathrm{SR}\sqrt{1 + \mathrm{b}_3 \frac{\mathrm{S}}{3}\mathrm{SR}} \tag{4}$$

Note that ASSR is the "adjusted for skewness Sharpe ratio," where SR is the Sharpe ratio, S is skewness, and b_3 is a parameter that expresses the investors' preference for skewness. The square root term can be seen as a multiplicative adjustment to the traditional Sharpe ratio. This adjustment depends upon the utility function of investors. For investors uninterested in skewness, say those with a quadratic utility function, b_3 is equal to zero and (4) collapses to the Sharpe ratio. In our empirical work, we explore two cases where b_3 takes on the values of 1 and 2. These two cases correspond to investors who have constant absolute risk aversion CARA and constant relative risk aversion CRRA, respectively.

To complicate matters, Lee and Su (2011) note that not only skewness but also the fat tails of the distribution are important in forecasting value at risk (VAR). Favre and Galeano (2002) propose adjustments to the Sharpe ratio to account for non-normality in the return distribution through the use of the VAR methodology. By maximizing returns that are subject to a maximum loss constraint for a given confidence interval, they derive the corresponding "adjusted Sharpe ratio" (ASR) as follows:

$$\mathrm{ASR} = \left[\frac{\mathrm{R_p} - \mathrm{R_f}}{\mathrm{VAR}} \right] \tag{5}$$

Note that R_p is the portfolio return, R_f is the three month Treasury interest rate, and VAR is a measure of risk.[12] Therefore, ASR is a reward-to-risk ratio like the Sharpe ratio.

While the ASR still assumes that returns are normally distributed, (5) can be extended to include other moments of the distribution. This "modified Sharpe ratio" (MSR) that adjusts the ASR to include the impact of skewness and excessive kurtosis of the return is shown below.

[12] Strictly speaking, this is a "stripped down" version of VAR as usually $VAR = V_p \sigma z T^{0.05}$ where V_p is the value of the portfolio, z is the z-value of the required percentile of the standard normal distribution, and T is the time in days it takes to windup a position. However, Favre and Galeano (2002) show that the value of the portfolio V_p appears in the numerator of the adjusted Sharpe ratio and so cancels out when the full VAR expression is included in the denominator. Therefore we neglect the term V_p as it will cancel out anyway in the ASR. In addition to this issue, Favre and Galeano (2002 neglect the square root of T. In essence they assume that the position can be liquidated in one day so the daily earnings at risk are the same as the value at risk.

$$\mathrm{MSR} = \left[\frac{\mathrm{R_p} - \mathrm{R_f}}{\mathrm{MVAR}} \right] \tag{6}$$

Now, MVAR is measured as follows.

$$MVAR = \left(z + \frac{1}{6}(z^2 - 1)S + \frac{1}{24}(z^3 - 3z)K - \frac{1}{36}(2z^3 - 5z)S^2 \right)\sigma$$

Notice that MVAR is simply VAR with the original *z* replaced in the above equation with the terms in brackets. The terms S and K refer to skewness and excess kurtosis, where, if S and K are zero, MVAR becomes VAR.

It is important to note that all of the above performance measures are static. To properly compare a passive strategy to an active strategy, one needs to use the MPPM, which examines whether managers are able to exploit arbitrage opportunities through dynamic trading. As noted earlier, Ingersoll et al. (2007) observe that using static performance measures to evaluate dynamic strategies can be misleading as portfolio managers can manipulate their strategies, either deliberately or otherwise, in order to score well on a wide variety of static performance measures even though the manager has no private information.

The MPPM Θ is

$$\Theta \equiv \left[\frac{1}{(1-\rho)\Delta t} \ln\left(\frac{1}{T} \sum_{t=1}^{T} \left[(1 + r_t)/(1 + r_{ft})^{(1-\rho)} \right] \right) \right] \tag{7}$$

The parameter ρ is the measure of relative risk aversion that, according to Ingersoll et al. (2007), historically varies between 2 and 4. Later, we find that there is no difference in our results whether we use 2, 3, or 4. Like Ingersoll et al. (2007), therefore, we report our results using a ρ value of 3. Meanwhile, Δt is the time length between observations, T is the number of observations, and r_t and r_{ft} are the portfolio's and the risk-free asset's un-annualised holding period return respectively at time t. The MPPM Θ measures the certainty equivalent risk premium earned by a given strategy relative to some benchmark. Our benchmark is the passive bond BH strategy, so positive values represent risk-adjusted superior performance and negative values represent inferior performance relative to the bond BH strategy.

5 Performance

The panels in Table 5 report the performance of the four investment strategies in order of ranking by the Sharpe ratio. The first panel represents performance for the full sample of 100 replications, and the second and third panels represent performance for the two equal sized sub-periods. This table also includes ranking by the MSR ratio as described by (6) and by the MPPM as described by (7).

<<Table 5 about here>>

Looking at performance from the static mean variance perspective, the best performing bond strategy is the forward curve strategy. This is distantly followed by the bond BH, immunization, and the slope premium strategies. A comparison of the Sharpe and MSR ratios suggests that adjusting for skewness and kurtosis does not adjust our perception of this performance since the MSR ratio ranks the four strategies in the same order as the Sharpe ratio. The picture is radically different when we examine the rankings according to the MPPM. Here, the passive bond BH strategy dominates all other strategies and the forward curve strategy, which was first-ranked from a static perspective, is ranked third. Clearly, it is vital to account for the dynamic nature of active strategies. Once this is accounted for, the active forward curve and slope premium strategies are seen to fall short of the passive bond BH strategy.

These radically different outcomes between static and dynamic performance measures are not hard to understand. Looking again at Table 3, we see that the forward curve strategy has deviant distributional properties, displaying the lowest standard deviation, the highest attractive positive skewness, and the highest unattractive excessive kurtosis. It appears that the mean variance measures, the Sharpe and the ASR ratios, as well as the distribution adjusted ratios, the MSR, CARA, and CRRA ratios are all inadvertently manipulated into ranking the forward curve strategy the highest. This occurs despite the fact that its actual excess return is lower than that of the bond BH and the slope premium strategies. Meanwhile, the MPPM rankings appear more reliable, where a modest extra return of 3 basis points with the slope premium strategy is not enough, on a risk-adjusted basis, to rank its performance above the passive bond BH strategy. Meanwhile, the attractive distributional properties (other than the mean) of the forward curve strategy are not enough to rank it ahead of the much more lucrative Bond BH strategy.

We check on the robustness of these conclusions by examining the performance of the four strategies for two sub-periods of equal size. These results, which are reported in the second and third panels of Table 5, show that the last sub-period agrees with the overall results but that the results of the first sub-period are different. Using the static MSR measure of performance, the ranking of the slope premium and the bond BH strategies switch places in the first period. This suggests that the MSR ratio provides additional information to investors concerned about excessive kurtosis. Even accounting for this, the forward curve strategy outperforms all other strategies across both sub-periods according to any measure of static performance. When measuring performance according to the dynamic MPPM, however, we see that the slope premium strategy performs best in the first period and, in a clear contrast to the static performance measures, the forward curve strategy performs the worst of all. It is interesting to note that the MPPM provides evidence that is consistent with findings from prior literature. Ilmanen (1995, 1997) finds that strategies like the slope premium strategy that use information in the term structure can form a viable bond strategy that can outperform a passive bond buy and hold strategy. Our results confirm this finding during a similar time period. However, markets are liable to trade away arbitrage opportunities and the results for the latter sub-period is consistent with this market behavior.

6 Simulation

To check on the robustness of our measures of mean, variance, skewness, and kurtosis, we explore the underlying distribution of these statistics by conducting a “bootstrap” simulation (see Davison and Hinkley 1997). Bootstrap experiments have become increasingly popular in the finance literature. For a good example see Chen, Huang and Lai (2011). From the original 100 x 4 strategy returns, we randomly select with replacement 100 rows of the returns. One row is selected each time so that the return period is the same for each strategy. We compute the mean, standard deviation, skewness, and kurtosis for each of these four strategies. This process is repeated 1,000 times. If our original statistics are reliable, then the 1,000 replications of these statistics should not be radically different and should preserve the rankings of the strategies by the size of a given statistic. That is, the strategy that has the highest (or lowest) skewness, for example, should consistently have the highest (or lowest) skewness throughout the experiment.

<Figures III and IV about here>

Figures III and IV plot the distribution of the standard deviation and skewness by percentiles for each of the four investment strategies. Comparing Figures III and IV to the corresponding values reported in Table 3, we see that the original statistics in Table 3 are about in the middle of the wide distribution of values obtained via the bootstrap. Moreover, the rankings of the strategies are consistent throughout the percentile range. For example, Figure IV agrees with Table 3 in that both show that the forward curve strategy consistently has the highest skewness of all four strategies throughout all percentiles.[13]

One important question to ask is whether the performance of the active strategies is significantly superior to the passive bond BH strategy and *vice versa*. Therefore, in addition to the descriptive statistics that are calculated for each simulated run, we also calculate the reported performance measures for each strategy as well as the log ratio of the performance measures of the active strategies relative to the bond BH strategy. Using the log ratio, if the performance measure of an active strategy is greater (or less) than the bond BH, the log ratio will be greater (or less)

[13] We also compile similar figures for mean and kurtosis and reach the same conclusion. Specifically, even the most extreme, 100th percentile mean and kurtosis are not very different from the corresponding statistics as reported for the original data in Table 3 and the rankings of the strategies by the size of mean and kurtosis are generally preserved throughout the experiment. Therefore, our mean and kurtosis statistics are also robust. These graphs are omitted for the sake of brevity and are available from the corresponding author upon request.

than zero. To test whether the performance measure of a candidate active strategy is significantly different, we examine the proportion of times that the bootstrapped log ratios are greater or less than zero. If the proportion of times the bootstrapped log ratio is greater than zero exceeds 97.5% (99.5%), then the candidate active strategy outperforms the bond BH at the 5% (1%) level of significance. On the other hand, if the proportion of times the bootstrapped log ratio is greater than zero falls below 2.5% (0.5%), then the bond BH outperforms the candidate active strategy at the 5% (1%) level of significance.

<<Table 6 about here>>

Table 6 reports the results of the log ratio tests of the performance ratios of a candidate strategy versus the passive bond BH strategy. Table 6 clearly shows that the forward curve possesses a statistically significant superior performance according to all the static measures of performance for the overall period. Meanwhile, the immunization and slope premium strategies typically perform worse than the bond BH strategy. These differences are usually statistically significant. Once we examine the dynamic MPPM, however, we find that the forward curve strategy is in fact significantly inferior to the bond BH strategy. Interestingly, using the MPPM performance measure, we are unable to find any statistical evidence that the bond BH strategy outperforms the active slope premium strategy.

Looking at the corresponding sub-period tests, we again note the differences in results according to the static and the dynamic performance measures. Specifically, the immunization and slope premium strategies continue to typically perform worse than the bond BH strategy while the forward curve strategy performs better according to the static measures of performance. However, these differences are often not statistically significant. This is particularly so for the forward curve strategy. The dynamic MPPM measure tests provide us with different results. These results confirm the story that is tentatively proposed by the MPPM performance measure in Table 5. For the first half sub-period, the passive bond BH strategy is significantly superior to all strategies except the slope premium strategy, whose performance is significantly superior to that of the bond BH strategy. In the second half sub-period, however, the passive bond BH strategy is significantly superior to all strategies including the slope premium strategy.

7 Conclusions

The most striking finding in this work is that the recently developed MPPM is much more critical for evaluating the performance of active strategies than are static measures of performance. When using the MPPM, we find that, consistent with prior literature, an active strategy based on time-varying term premiums can indeed form the basis of a successful bond strategy that outperforms an unbiased expectation-inspired passive bond buy and hold strategy. As we stated at the outset, however, this is only true of the earlier time period when the literature first made this claim. For a later time period, we find that the passive buy and hold strategy is significantly superior to all active strategies. This result is confirmed by statistical tests, and it suggests that once it became known that an active strategy based on time-varying term premiums could outperform a passive buy and hold strategy, the markets adjusted and arbitraged away this opportunity.

In contrast to the MPPM, we find that all static measures of performance are unable to detect this change. They rather uncritically rank the deviant forward curve strategy as the best both overall and in all sub-periods in a statistically significant way, even though the mean return of this strategy is 28 basis points lower than that of the passive bond buy and hold strategy. Consequently, we find that the dynamic MPPM is more consistent with the underlying data, and we recommend that one should rely on the MPPM in order to measure the performance of dynamic strategies.

Overall, it appears that the unbiased expectation hypothesis is the more likely explanation of the behaviour of the term structure in recent times, as economically and statistically significant superior performance cannot be achieved if one uses information in the forward curve or the term structure as a guide to adjusting bond portfolios in response to changes in the term premium.

Acknowledgements

This work was supported by Junta de Comunidades de Castilla-La Mancha [grant number PEII11-0031-6939] and Ministerio de Ciencia e Innovación [grant number ECO2011-28134], and it was partially supported by Fondo Europeo de Desarrollo Regional (FEDER) funds.

References

Ang A, Piazzesi M and Wei M (2006) What does the yield curve tell us about GDP growth? J Econom 131:359-403.

Bierwag G O (1979) Dynamic portfolio immunization policies. J Bank. Finance 3:23-41.

Chen C. –W, Huang C. –S and Lai H. –W (2011) Data snooping on technical analysis: evidence from the Taiwan stock market. Rev Pac Basin Financial Mark Policies 14: 195-212.

Cochrane J, and Piazzesi M (2005) Bond risk premia. Am Econ Rev 95:138-160.

Davison A, and Hinkley D (1997) Bootstrap methods and their application. Cambridge University Press, Cambridge.

De Bondt W, and Bange M (1992) Inflation forecast errors and time variation in term premia. J Financial Quant Anal 27:479-496.

Diaz A, Navarro E, Gonzales M, and Skinner F (2009) An evaluation of contingent immunization. J Bank Finance 33:1874-1883.

Estrella A, and Mishkin F (1997) The predictive power of the term structure of interest rates in Europe and the United States: implications for the European Central Bank. Eur Econ Rev 41:1375-1401.

Estrella A, and Mishkin F (1998) Predicting U.S. recessions: financial variables as leading indicators. Rev Econ Stat 80:45-61.

Fama E, and Bliss R (1987) The information in long-maturity forward rates. Am Econ Rev 77:680-692.

Favre L, and Galeano J (2002) Mean modified value-at-risk optimization with hedge funds. J Altern Invest 5:21-25.

Fong G, and Vasicek O (1984) A risk minimizing strategy for multiple liability immunization. J Finance 39:1541-1546.

Froot K A (1989) New hope for the expectations hypothesis of the term structure of interest rates. J Finance 44:283-306.

Galvani V, and Landon S (2011) Riding the yield curve: a spanning analysis. Rev Quant Finance and Acc (forthcoming)

Gregoriou G, and Gueyie J (2003) Risk adjusted performance of funds of hedge funds using a modified Sharpe ratio. J Altern Invest 6:5-8.

Hardouvelis G. (1988) The predictive power of the term structure during recent monetary regimes. J Finance 43:339-356.

Henriksson R, and Merton R (1981) On market timing and investment performance II: Statistical procedures for evaluating forecasting skills. J Bus 54:513-534.

Ilmanen I., (1995) Time-varying expected returns in international bond markets. J Finance 50, 481-506.

Ilmanen I., (1996) Market expectations and forward rates. J Fixed Income 6, 8-22.

Ilmanen I., (1997) Forecasting U.S. bond returns. J Fixed Income 7, 22-37.

Ilmanen I., Sayood R (2002) Quantitative forecasting models and active diversification for international bonds. J Fixed Income 12, 40-51.

Ingersoll J., Spiegel M., Goetzmann W., and Welch I (2007) Portfolio performance manipulation and manipulation-proof performance measures. Rev Financial Stud 20, 1503-1546.

Kalev P, and Inder B (2006) The information content of the term structure of interest rates. Appl Econ 38, 33-45.

Kessler S, and Scherer B (2009) Varying risk premia in international bond markets. J Bank Finance 33:1361-1375.

Lee C. –F and Su J. –B (2011) Alternative statistical distributions for estimating value at risk: theory and evidence. Rev Quant Finance and Acc (forthcoming)

Lehmann E (1975) Nonparametrics: statistical methods based on ranks. Holden-Day, San Francisco.

Longstaff F (1990) Time varying term premia and traditional hypotheses about the term structure. J Finance 45:1307-1314.

Longstaff F (2000) Arbitrage and the expectations hypothesis. J Finance 55:989-994.

Henriksson R, and Merton R (1981) On market timing and investment performance II: Statistical procedures for evaluating forecasting skills. J Bus 54:513-534.

Modigliani M, and Sutch R (1966) Innovations In Interest Rate Policy. Am Econ Rev 56:178-197.

Newey W, and West K (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55:703-708

Papageorgiou N, and Skinner S (2002) Predicting the Direction of Interest Rate Movements. J Fixed Income 11: 87-95.

Soto G (2001) Immunization derived from a polynomial duration vector in the Spanish bond market. J Bank Finance 25:1037-1057.

Van Horne J C (1980) The term structure: A test of the segmented markets hypothesis. South Econ J 46:1129-1140.

Ventura J, and Pereira C (2006) Immunization using a stochastic-process independent multi-factor model: The Portuguese experience. J Bank Finance 30:133-156.

Zakamouline V, and Koekebakker S (2009) Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. J Bank Finance 33:1242-1254.

Table 1

Statistical test of term premium forecast success rates

	Increase	*No Increase*	*Total*
Predict	1,842 (n)	2,494 (n)	4,336
Success	936 (n_1)	1,525 (n_1)	2,461
Fail to Predict	970	905	1,87
Total	1,906 (N_1)	2,430 (N_1)	4,336 (N)
Expected Random Success (nxN_1/N)	810	1,398	2,208
Excess Success	126***	127***	253

Note: This table shows the predictions for both increase and no increase in the ten-year rate of interest. ***Statistically significant at the 1% level.

Table 2

Statistical test of the forward curve prediction of future bond returns

X	*Mo 1*	*Mo 3*	*Mo 6*	*Mo 9*	*Mo 12*
γ_0	-0.009	-0.013	0.000	0.002	0.008
	(-1.445)	(-1.794)	(0.029)	(0.244)	(1.054)
γ_1	-2.733	-1.464	0.695	1.814	3.721
	(-10.841)	(-5.247)	(2.564)	(6.783)	(14.640)
γ_2	-0.737	-2.733	-7.040	-7.627	-6.857
	(-1.064)	(-3.644)	(-8.611)	(-8.930)	(-8.211)
γ_3	8.520	8.854	9.573	9.776	4.050
	(6.043)	(6.942)	(7.023)	(6.856)	(2.873)
γ_4	-1.954	-1.863	3.223	2.054	1.200
	(-1.325)	(-1.490)	(2.535)	(1.558)	(0.878)
γ_5	-0.913	-0.469	-1.040	-0.237	1.664
	(-1.292)	(-0.565)	(-1.200)	(-0.235)	(1.450)
γ_6	8.892	7.199	1.449	-2.788	-2.108
	(8.568)	(5.924)	(1.147)	(-2.338)	(-1.903)
γ_7	-10.506	-8.869	-6.270	-2.325	-1.019
	(-13.934)	(-9.438)	(-6.379)	(-2.603)	(-1.158)
R^2	0.457	0.339	0.232	0.167	0.167

Note: This table shows the results of the full period regression of one-year excess returns on the ten-year benchmark Treasury bond *x* months in the future on the interest rate swap forward curve. Specifically, the regression is $R_{t+x}^N = \gamma_0 + \gamma_1 S_{1,t} + \gamma_2 F_{1,t} + \gamma_3 F_{2,t} + \gamma_4 F_{3,t} + \gamma_5 F_{4,t} + \gamma_6 F_{6,t} + \gamma_7 F_{9,t} + \varepsilon_t$, where R_{t+x}^N is the one-year excess (above the three-month t-bill rate) total return on the benchmark ten-tear Treasury bond *x* months in the future, S_1 is the one year swap interest rate, and F_2, F_3 F_4, F_6, F_9 are the two-, three-, four-, six-, and nine-year forward rates, respectively. T-statistics are in parentheses.

Table 3

Statistical characteristics of the holding period returns

	Mean	*SD*	*Skewness*	*Excess Kurtosis*	*Range*	*Minimum*	*Maximum*
Full Period							
Bond BH	0.0721	0.0101	0.7112	-0.0569	0.0451	0.0572	0.1022
Forward Curve	0.0693	0.0079	0.9543	0.4002	0.0362	0.0568	0.0930
Slope Premium	0.0724	0.0113	0.5806	-0.1405	0.0514	0.0534	0.1048
Immunization	0.0688	0.0091	0.5190	-0.7110	0.0353	0.0545	0.0898
First Half							
Bond BH	0.0780	0.0100	0.2907	-0.5639	0.0399	0.0623	0.1022
Forward Curve	0.0737	0.0084	0.3671	-0.7539	0.0316	0.0614	0.0930
Slope Premium	0.0791	0.0106	0.3273	-0.6064	0.0420	0.0628	0.1048
Immunization	0.0736	0.0093	-0.0771	-1.1730	0.0322	0.0576	0.0898
Second Half							
Bond BH	0.0663	0.0061	0.4431	-0.9301	0.0216	0.0572	0.0787
Forward Curve	0.0649	0.0040	0.0441	-0.2116	0.0166	0.0568	0.0734
Slope Premium	0.0656	0.0073	0.2253	-0.9439	0.0261	0.0534	0.0795
Immunization	0.0639	0.0058	0.3686	-0.7998	0.0215	0.0545	0.0761

Note: This table reports the characteristics of the distribution of 100 nine-year annualized holding period returns starting from January 2, 1990 for four long-term investment strategies. This information is repeated for the first and second half sub-periods from January 2, 1990 to February 1 1994 and from March 1, 1994 to April 2, 2007, respectively.

Table 4
Statistical significance of alternative strategies versus the buy and hold strategy

	Mean Difference in %	*Newey West SE*
Full Period		
Forward Curve-Bond BH	-0.2884***	0.1075
Immunization-Bond BH	-0.3356***	0.0551
Slope Premium-Bond BH	0.0201	0.6686
First half		
Forward Curve-Bond BH	-0.4255***	0.1206
Immunization-Bond BH	-0.4358***	0.0795
Slope Premium-Bond BH	0.1187**	0.0488
Second half		
Forward Curve-Bond BH	-0.1512	0.1743
Immunization-Bond BH	-0.2354***	0.0206
Slope Premium-Bond BH	-0.0784	0.4187

Note: This table reports whether the forward curve, immunization, and slope premium strategies achieve a return that is significantly different from the return of the passive bond BH strategy. Newey West SE are the autocorrelation- and heteroskedasticity-corrected regression standard errors, which are found by regressing the differences in the return of the bond BH and the candidate strategy on a constant and are used to test whether the reported means of the candidate strategy are significantly different from those of the bond BH strategy. ***, **significant at the 1% and 5% levels, respectively.

Table 5
Performance of the strategies

Full Period	*MSR Rank*	*MPPM Rank*	*Total Return*	*Sharpe Ratio*	*CARA*	*CRRA*	*ASR*	*MSR*	*MPPM*
Full Period									
F. Curve	1	3	6.929	3.472	5.217	6.510	1.490	2.521	-15.501
Bond BH	2	1	7.212	2.992	4.342	5.362	1.284	1.870	0.000
Immunization	3	4	6.883	2.978	4.304	5.308	1.278	1.784	-18.679
Slope Premium	4	2	7.239	2.699	3.812	4.667	1.159	1.551	-0.244
First Half									
F. Curve	1	4	7.369	3.524	5.317	6.643	1.513	1.924	-24.159
Bond BH	3	2	7.796	3.397	5.102	6.366	1.458	1.741	0.000
Slope Premium	2	1	7.914	3.302	4.916	6.118	1.417	1.733	6.692
Immunization	4	3	7.360	3.162	4.644	5.756	1.357	1.500	-23.972
Second Half									
F. Curve	1	3	6.488	6.315	11.449	14.909	2.710	2.808	-7.311
Bond BH	2	1	6.643	4.458	7.256	9.242	1.913	2.614	0.000
Immunization	3	4	6.407	4.286	6.866	8.714	1.839	2.332	-13.700
Slope Premium	4	2	6.564	3.579	5.449	6.825	1.536	1.862	-6.629

Note: The full period reports the annualized nine-year holding period return and the performance for the full 100-month period from January 4, 1999 to April 30, 2007. The first and second half periods report the same for January 4, 1999 to February 3, 2003 and March 3, 2003 to April 30, 2007, respectively. The Sharpe ratio is the traditional mean variance performance measure. CRRA are Sharpe ratios adjusted for skewness for investors who have a constant relative risk aversion. ASR is the Sharpe ratio adjusted for tail risk, and MSR is the Sharpe ratio adjusted for skewness and kurtosis as well as tail risk. Finally, MPPM is the manipulation-proof performance measure that is designed to compare the performance of a given strategy to the bond buy and hold bond BH strategy for investors with relative risk aversion.

Table 6

Log-ratio test of the statistical significance of alternative strategies versus the buy and hold strategy

Percentile	*MSR Rank*	*MPPM Rank*	*Total Return*	*Sharpe Ratio*	*CARA*	*CRRA*	*ASR*	*MSR*	*MPPM*
Full Period			Dif in TR	LR Test Statistic	LR Test Statistic	LR Test Statistic	LR Test Statistic	LR Test Statistic	LR Test Statistic
vs F. Curve	1	3	-0.0029	0.140***	0.235***	0.270***	0.140***	0.234***	-0.089***
Prob(LR>0)				[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.000]
vs Immunization	3	4	-0.0034	-0.015	-0.080***	-0.108***	-0.015	-0.084***	-0.108***
				[0.232]	[0.000]	[0.000]	[0.232]	[0.002]	[0.000]
vs Slope Premium	4	2	0.0002	-0.114***	-0.179***	-0.208***	-0.114***	-0.180***	-0.001
				[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.440]
First Half									
Vs F. Curve	1	4	-0.0043	0.037	0.073	0.094	0.037	0.080	-0.127***
				[0.786]	[0.832]	[0.826]	[0.786]	[0.881]	[0.000]
vs Immunization	4	3	-0.0044	-0.071**	-0.251***	-0.404**	-0.071**	-0.131***	-0.126***
				[0.004]	[0.000]	[0.020]	[0.004]	[0.000]	[0.000]
vs Slope Premium	3	1	0.0012	-0.028***	-0.017	0.010	-0.028***	-0.012**	0.033***
				[0.000]	[0.126]	[0.360]	[0.000]	[0.030]	[1.000]
Second Half									
vs F. Curve	1	3	-0.0015	0.351***	0.141	0.014	0.351***	0.159	-0.046***
				[1.000]	[0.771]	[0.513]	[1.000]	[0.982]	[0.001]
vs Immunization	3	4	-0.0024	-0.040**	-0.090***	-0.113***	-0.040***	-0.090***	-0.088***
				[0.012]	[0.003]	[0.003]	[0.012]	[0.001]	[0.000]
vs Slope Premium	4	2	-0.0008	-0.220***	-0.356***	-0.429***	-0.220***	-0.303***	-0.042***
				[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Note: This table reports the log-ratio test of the significance of differences in the performance of the passive bond BH as compared to the forward curve, immunization, and slope premium strategies. These tests are reported for the overall period and for the first half and second half of the full period, which are January 4, 1999 to February 3, 2003 and March 3, 2003 to April 30, 2007, respectively. The Sharpe ratio is the traditional mean variance performance measure. CRRA are Sharpe ratios adjusted for skewness for investors who have a constant relative risk aversion. ASR is the Sharpe ratio adjusted for tail risk, and MSR is the Sharpe ratio adjusted for skewness and kurtosis as well as tail risk. Finally, MPPM is the manipulation-proof performance measure that is designed to compare the performance of a given strategy to the bond buy and hold bond BH strategy for investors with relative risk aversion. ***, **significant at the 1% and 5% levels, respectively.

Figure I This figure plots the ten year and one year Treasury yields and the spread between them from January 2, 1990 to April 30, 2007

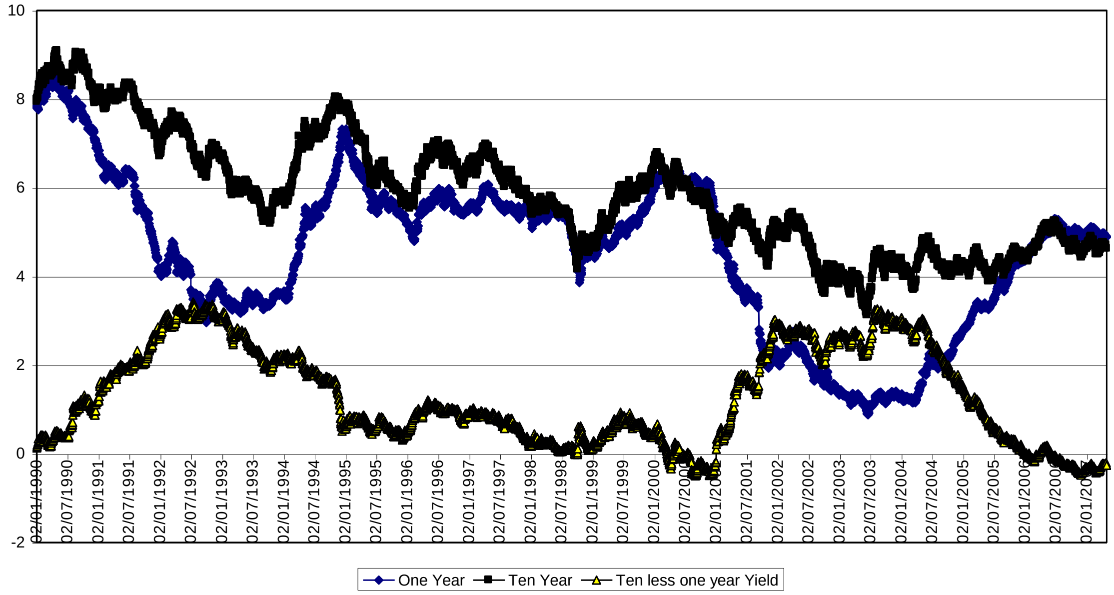

Figure II: This figure plots the excess holding period returns for four bond strategies, the Bond BH, Slope Premium, Forward Curve and the Immunization strategies.

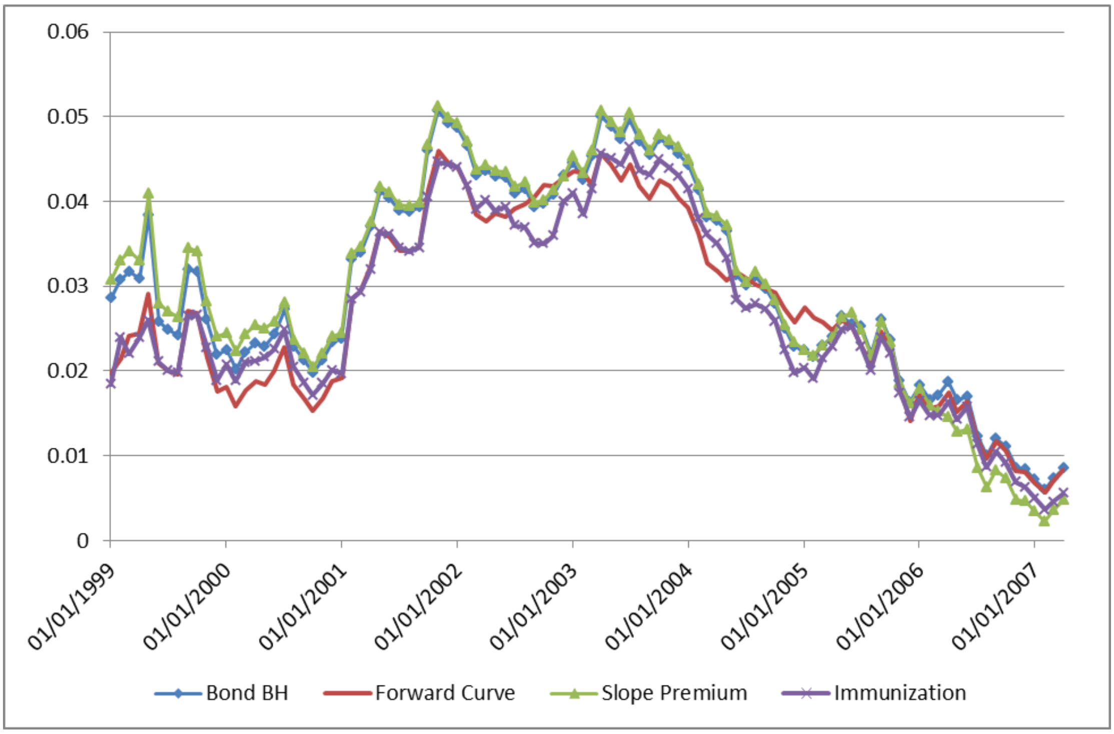

Figure III: This figure represents the bootstrapped standard deviation sorted by strategy as indicated in the legend. We calculate the bootstrapped SD by resampling from the original holding period returns 1,000 times and plot the results by percentiles.

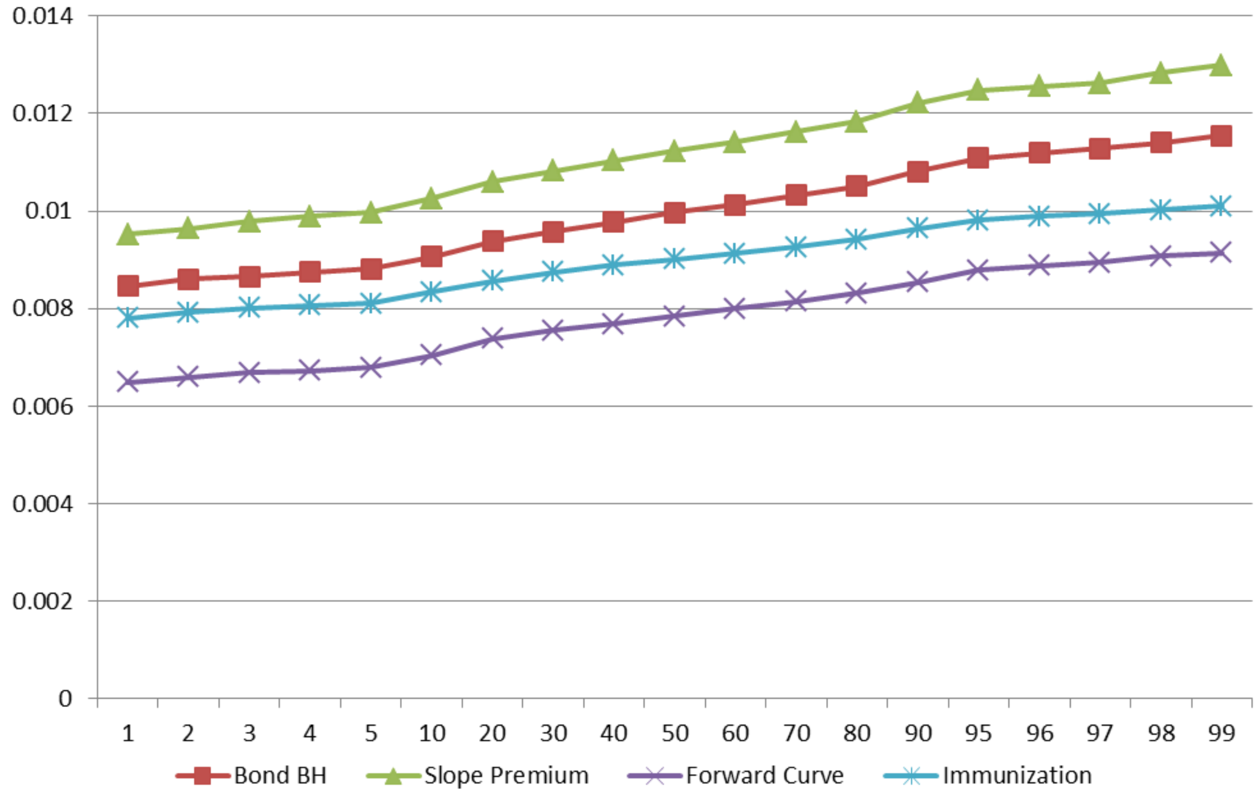

Figure IV: This figure represents the bootstrapped skewness sorted by strategy as indicated in the legend. We calculate the bootstrapped skewness by resampling from the original holding period returns 1,000 times and plot the results by percentiles.