NEAREST-NEIGHBOR FORECASTS OF U.S. INTEREST RATES

JOHN BARKOULAS
University of Tennessee

CHRISTOPHER F. BAUM
Boston College

ATREYA CHAKRABORTY
Cambridge, MA 02144

Abstract

We employ a nonlinear, nonparametric method to model the stochastic behavior of changes in several short and long term U.S. interest rates. We apply a nonlinear autoregression to the series using the locally weighted regression (LWR) estimation method, a nearest-neighbor method, and evaluate the forecasting performance with a measure of root mean square error (RMSE). We compare the forecasting performance of the nonparametric fit to the performance of two benchmark linear models: an autoregressive model and a random-walk-with-drift model. The nonparametric model exhibits greater out-of-sample forecast accuracy that that of the linear predictors for most U.S. interest rate series. The improvements in forecast accuracy are statistically significant and robust. This evidence establishes the presence of significant nonlinear mean predictability in U.S. interest rates, as well as the usefulness of the LWR method as as modeling strategy for these benchmark series.

1. Introduction

There has been considerable interest in modeling the evolution and interactions of nominal interest rates. Models of the term structure of interest rates fall into two categories: those based upon arbitrage arguments and those based on a general equilibrium formulation. In the former category, single-factor models of the term structure of interest rates have been proposed by Merton (1973), Vasicek (1977), Dothan

(1978), Schaefer and Schwartz (1978), and many others. Multifactor term structure models have been proposed by Richard (1987), Brennan and Schwartz (1979), Langetieg (1980), Schaefer and Schwartz (1984), and Heath et al. (1992). Models representing a complete general equilibrium specification of the term structure have been put forth by Cox, Ingersoll and Ross (1985a,b), Longstaff and Schwartz (1992) and many others. Chan, Karolyi, Longstaff and Sanders (1992) and Broze, Scaillet, and Zakoian (1995) provided empirical comparisons of the adequacy of the models' explanation of the data. Both theoretical and empirical results from these two branches of research suggest that term structure models that allow yield nonlinearity can provide additional insight and explanatory power for the modelling of equilibrium interest rates.

This paper models nonlinearities in the evolution of the conditional means of U.S. Treasury securities' yield changes for various maturities and attempts to exploit those nonlinearities to improve forecasting performance. Our approach is nonstructural and univariate as it relies on the historical behavior of individual securities' yield series to model nonlinearities. It is also nonparametric as we do not impose maintained hypotheses of smoothness on the regression function; instead, we let the data determine the regression function. To detect whether nonlinearity characterizes the securities' yields, we apply the BDS test to the series. We also investigate the low-frequency properties of securities yields by subjecting the series to unit root tests that consider both integer and fractional orders of integration.

We are primarily interested in the out-of-sample forecasting accuracy of our nonparametric model as measured by root mean square error (RMSE). We employ the locally weighted regression (LWR) estimation method, a local-fitting methodology, to provide a regression surface and carry out prediction. We compare the forecasting performance of the nonparametric fit to the performance of two benchmark linear models: an autoregressive (AR) model and a random-walk-with-drift (RW) model. The overall evidence shows improved forecastibility with nonparametric predictors out of sample over best linear predictors for most U.S. interest rate series.

The plan of the paper is as follows. In section 2 we present the nonparametric method. Data and diagnostic tests are presented in section 3, while the empirical estimates are reported in section 4. Finally, we conclude in section 5 with a summary of our results and suggestions for future research.

2. Econometric Methodology

We attempt to uncover nonlinear relationships in U.S. T-bill and bond yields using the nonparametric locally weighted regression (LWR) method. LWR is a nearest-neighbor estimation technique, first introduced by Cleveland (1979) and further developed by Cleveland and Devlin (1988), and Cleveland, Devlin, and Grosse (1988). It is a way of estimating a regression surface through a multivariate smoothing procedure, fitting a function of independent variables locally and in a moving-average manner.

Suppose that the regression function is given by

$$y_t = g(x_t) + \varepsilon_t, \qquad t = 1, \ldots, n \tag{1}$$

where $x_t = (x_{1t}, \ldots, x_{pt})$ is a $1 \times p$ vector of (weakly) exogenous explanatory variables, $g(\cdot)$ is a smooth function mapping $R^p \to R$, and ε_t is an independent and identically distributed disturbance with mean zero and variance σ^2.

LWR is a numerical algorithm that describes how $\hat{g}(x^*)$ the estimate of g at the specific value of x^*, is estimated. Let f be a smoothing constant such that $0 < f \leq 1$, and let $q_f = \text{int}(f.n)$, where $\text{int}(\cdot)$ extracts the integer part of its argument. The LWR uses the "window" of q_f observations nearest x^*, where proximity is defined using the Euclidean distance. In the LWR algorithm, the conditional mean is estimated from a weighted least squares regression of y on x for the relevant q_f observations. More specifically, given a point x^* called the current state, rank the x_t's by Euclidean distance from x^*. Let $\|\cdot\|$ measure Euclidean distance; then the Euclidean distance from x^* to its x_{q_f} nearest neighbors is

$$\mathrm{d}\left(x^*, x_{q_f}\right) = \left[\sum_{j=1}^{p}\left(x_{q_{fj}} - x_j^*\right)\right]^{1/2} \tag{2}$$

Each of the q_f nearest neighbors are inversely weighted by their Euclidean distance from the current state. Let $w_{it} = 1\text{-}u$, where

$$u \equiv \frac{\left\|x_{it} - x_t^*\right\|}{\sum_{i=1}^{q}\left\|x_{it} - x_t^*\right\|} \tag{3}$$

The remaining observations are assigned a weight of zero. We also tried the tricube function, $w_{ti} = (1-u^3)^3$, suggested by Cleveland, as well as locally unweighted regression but (3) proved slightly superior empirically.

The value of the regression surface x^* at is then computed as

$$\hat{y}^* = \hat{g}(x^*) = x^{*\prime}\hat{\beta}, \tag{4}$$

where

$$\hat{\beta} = \arg\min\left[\sum_{t=1}^{n} w_t\left(y_t - x_t{}'\beta\right)^2\right]. \tag{5}$$

Stone (1977) addressed the issue of consistent estimation through regularity conditions on weights of the neighbors. Consistency of NN estimators (and therefore LWR) requires that the number of NNs used go to infinity with sample size, but at a slower rate, that is as $n \to \infty, q \to \infty$, but $q/n \to 0$. Consistency becomes a matter of imposing a selection rule on f. As f increases the number of NNs (neighborhood size) increases, the bias in $\hat{g}(x)$ tends to increase, and the sampling variability tends to decrease. In practice one needs to choose to balance the trade-off between bias and variance.

The LWR estimator of g(.) is linear in y :

$$\hat{g}(x^*) = \sum_{i=1}^{n} l_i(x^*) y_i \tag{6}$$

where the $l_i(x)$ depend on x_t, $t = 1, \ldots, n, w, d$, and f, but not on the y_t. Therefore the statistical properties of the estimators can be derived with standard techniques. A difficulty arises since the projection $(I - L)$ matrix which delivers LWR residuals is neither idempotent nor symmetric. Although the exact distribution of the error sum of squares is not χ^2 (as the eigenvalues of $(I - L)$ need not be all ones or zeros), it can be approximated by a constant multiplied by a χ^2 variable. The constant and degrees of freedom are chosen so that the first two moments of the approximating distribution match those of the distribution of the error sum of squares (Kendall and Stuart, 1977).

3. Data and Diagnostic Tests

Our data set consists of a variety of short- and long-term U.S. Treasury interest rates: monthly observations on the Federal Funds rate, 3-month, 6-month, and 12-month U.S. Treasury bill yields, and yields on 5-year and 10-year Treasury bonds. The sample period is 1957:01 to 1993:12 for all series except for the 6-month Treasury-bill yield, for which it is 1959:01 to 1993:12. The 3- and 6-month yields are percentage annual rates obtained from the secondary market. The 1-, 5-, and 10-year yields are constant maturity percentage annual rates. All data series are obtained from the Citibank data base. As we are primarily interested in the out-of-sample forecasting performance of the nonparametric method, we reserve the last 48 observations (1989:01 to 1993:12) from each yield-change series for forecasting purposes. Diagnostic tests presented below are applied to the estimation sample series, excluding the post-sample observations.

Table 1 presents selected summary statistics for first-differenced yield series. The means for all series are not statistically different from zero and all series exhibit dependence in the third and fourth cumulants. All yield-change series are negatively skewed with the exception of the 5-year yield which is positively skewed. They are all characterized by fat tails, i. e., leptokurtosis. However, the presence of skeweness and kurtosis is much stronger for the short term as opposed to longer term interest rates. As expected, yield changes for short maturities exhibit greater variability than those for longer maturities.

We first investigate the low-frequency properties of the yield series. To do so, we apply the Phillips-Perron tests (PP) (Phillips (1987), Phillips and Perron (1988)) to both levels and first differences of our sample series, with results presented in Table 2. Inference is robust to the order of serial correlation allowed in the data. All PP tests fail to reject the unit root null hypothesis in the yield series but strongly reject the unit root null in yield changes. The PP test results therefore strongly support the hypothesis of a single unit root in the yield series. Given the low power of standard unit root tests against fractional alternatives (Diebold and Rudebusch (1991)) we apply the semi-nonparametric procedure suggested by Geweke and Porter-Hudak (GPH, 1983) to the yield series. The GPH test avoids the knife-edged $I(1)$ and $I(0)$ distinction in the PP test by allowing the integration order to take on any real value (fractional integration). Table 3 reports the empirical estimates for the fractional differencing parameter. We find no evidence in support of the fractional alternative for any of our sample series. We therefore conclude that all yield series are integrated processes of order one and subsequently apply our analysis to yield changes.[1]

To obtain some preliminary evidence regarding the presence of nonlinearities, we perform the test suggested by Brock, Dechert, and Scheinckman (BDS, 1987) to yield changes and filtered yield changes. The BDS test tests the null hypothesis of independent and identical distribution (i.i.d.) in the data against an unspecified departure from i.i.d. A rejection of the i.i.d. null hypothesis in the BDS test is consistent with some type of dependence in the data, which could result from a linear stochastic system, a nonlinear stochastic system, or a nonlinear deterministic system. Under the null hypothesis, the BDS test statistic asymptotically converges to a standard normal variate. However, Monte Carlo simulations by Brock, Hsieh, and LeBaron (1991) and Hsieh and LeBaron (1988) suggest that the asymptotic distribution is a poor approximation to the finite sample distribution when there are fewer than 500 observations, and is not appropriate when applied to the standardized residuals of ARCH models.[2]

Table 4 reports the BDS test statistics for three sets of data: the yield-change series and two prewhitened versions created with autoregressive and autoregressive conditionally heteroscedastic (ARCH) model filters. Going from the shorter to longer maturity yield series, the AR orders chosen on the basis of AIC are 13, 20, 19, 19, 6, and 22, respectively (the maximum order allowed is 24), and the ARCH orders are 3, 4, 2, 2, 2, and 4, respectively.[3] We applied the BDS test to these three sets of series for embedding dimensions of $m = 2,3,4$ and 5. For each m, ε is set to 0.5 and 1.0 standard deviations (σ) of the data. We use the quantiles from the small sample simulations reported by Brock et al. (1991) as approximations to the finite-sample critical values of our BDS statistics. The i.i.d. null hypothesis is overwhelmingly rejected in all cases for yield changes. When the BDS test is applied to the AR-filtered series we still obtain strong rejections of the i.i.d. null hypothesis suggesting that linear dependence in the first moments does not fully account for rejection of i.i.d. in yield changes. When applied to the ARCH model, the obtained residuals are standardized by their estimated conditional standard deviations and the BDS test is applied to these standardized residual series. The standardized residuals appear to be i.i.d. suggesting no neglected nonlinearity in the series after properly accounting for time variation in their second moments. However, the estimated ARCH effects could be proxying nonlinearity in the conditional mean of our series.[4] As we are interested in uncovering nonlinear structure in the conditional mean of yield-change series, we now turn to modeling nonlinearities in their first moments by means of local-fitting methodology.

4. Empirical Results

In this section we estimate a nonlinear model for the yield-change series and compare its ex ante forecasting performance to that of benchmark linear models. In modeling nonlinearities in the conditional mean of our sample yield-change series we allow for a nonparametric functional form in the relationship. Employing nonparametric regression in the estimation process has the following advantages: (i) it reduces the possibility of model misspecification, (ii) it provides a versatile method of exploring a general relationship, and (iii) it can serve as a diagnostic tool in suggesting simple parametric formulations of the regression relationship. The nonparametric method employed is locally weighted regression (LWR), as specified in section 2 above.

We compare the LWR forecasts to those obtained by estimating two standard linear models: an autoregressive model (AR) and a random-walk-with-drift model (RW). The last 48 observations (1989:01 to 1993:12) from each yield-change series are reserved for forecasting purposes. The first twenty-four yield changes (1957:02 to 1959:02; 1959:02 to 1961:02 for the 6-month yield changes) are used for model initialization. Therefore, for forecasting purposes the sample period 1959:03 (1961:03 for the 6-month rate) to 1988:12 is the training set and the sample period 1989:01 to 1993:12 is the test set. The out-of-sample forecasting horizon is one-step ahead, with the competing forecasting models estimated over the training set and applied over the test set to generate genuine one-step-ahead out-of-sample forecasts. The criterion for forecasting performance is root mean square error (RMSE).[5] The AR order for the linear model is chosen on the basis of AIC, as described in section 3.

Tables 5 through 10 report the out-of-sample forecasting performance of the LWR model and its linear counterparts for our six sample series.[6] To ensure robustness of our evidence, we report the forecasting performance of nonlinear autoregressions of order one through six and for varying window sizes.[7] Comparing the forecasting performance of the linear models, we observe that the AR model outperforms the RW model for all yield series except for 3-month and 12-month maturities. We therefore restrict ourselves to comparisons of the forecasting performance between the nonparametric model and the AR model.

On the basis of the RMSE forecasting criterion, the LWR model outperforms the AR model across the different lag structures and window sizes considered. Only in a very limited number of cases is the AR fit better than the LWR fit and these primarily concentrate on the 5-year yield changes. The percentage reductions in RMSE of the LWR fit over the AR fit range from 1.08% to 15.54% for the Federal Funds rate, 6.79% to 20.34% for the 3-month maturity, 2.75% to 13.40% for the 6-month maturity, 7.71% to 21.14% for the 12-month maturity, 0.002% to 5.54% for the 5-year maturity, and 1.16% to 7.20% for the 10-year maturity. The improvements in forecasting performance appear to be greater for the short-term as opposed to longer-term yield changes. The consistency in the forecasting improvements of the nonparametric specification with respect to autoregression order and window size enhances the view that the estimated nonlinearities are not a statistical artifact, but rather capture essential aspects of the data generating process.

In order to formally evaluate model performance, we apply the forecast comparison test of Granger and Newbold (1986, pp. 278-280) to test the hypothesis that there is no difference in the forecasting accuracy between the linear and nonlinear models. Given that the nonparametric fit generally achieves a lower RMSE, we test the null hypothesis of no difference in the forecasting performance between the AR and LWR models against the one-sided alternative that the LWR model has superior forecasting performance. Table 11 reports the Granger-Newbold test results. For the 3-month, 6-month, 12-month and 10-year yield-change series, the forecasting improvements obtained by the nonparametric model are statistically significant. For the 5-year yield series, we cannot reject the null hypothesis of no difference in the forecasting accuracy of the two competing models. This is not surprising given that the RMSE values attained by the LWR method were not uniformly lower than those attained by the AR method, and in cases when they were lower, they were only marginally so. Finally, for the Federal Funds rate, the LWR method attains a lower RMSE in 51 out of 54 cases considered, but fails to demonstrate superiority in a statistically significant sense. Overall, the test results show that the LWR model's forecast performance is statistically superior to that of the AR model for most of our sample series. This enhances the view that the LWR method is a superior modeling strategy for these benchmark interest rate series.

5. Conclusions

We provide evidence that the nonparametric fit generated by the LWR method results in significant improvements, compared to benchmark (linear) models, in out-of-sample U.S. short- and long-term interest rate forecasts. The superior performance of the LWR methodology is robust to autoregression order and window size. The evidence provided establishes the presence of significant nonlinear mean predictability in U.S. interest rates, as suggested by theoretical findings. Since it appears that the dynamical behavior of US interest rates is very local in nature, letting the data determine the regression function pays dividends in terms of improvements in the forecasting accuracy of interest rate models. Although the LWR estimation method has failed to successfully predict stock returns (Hsieh (1991), LeBaron (1988)) and exchange rates (Diebold and Nason (1990), Meese and Rose (1990, 1991), Mizrach (1992)) it appears to be very useful in modeling conditional mean changes in interest rate series.[8] In addition to nonlinearities in variances and possibly higher moments, significant nonlinearities in the mean clearly exist for U.S. interest rates, and LWR appears to successfully capture those nonlinearities.

Our results can be extended in several ways. First, the usefulness of the LWR estimation method as a forecast generating mechanism for multiple-step-ahead forecasting horizons should be investigated. Second, our positive results invite the use of alternative nonparametric methods to be employed as forecasting tools for U.S. interest rates. Finally, an obvious avenue of future research is to apply the LWR methodology to interest rate series from other industrial countries.

Endnotes

1. We also applied the KPSS test (Kwiatkowski, Phillips, Schmidt, and Shin (1992)) in which the null hypothesis is stationarity. We are able to reject the trend-stationarity null for reasonable lag structures, thus suggesting the presence of a unit root in the series. These results are not reported here but are available upon request.
2. Brock's (1986) residual theorem, stating that the asymptotic distribution of the BDS test is not altered by using residuals instead of raw data in linear models, extends to some nonlinear models but not to ARCH models.
3. The orders for the conditional variance equation are chosen on the basis of superior performance of diagnostic tests for serial

correlation in the standardized and squared standardized residuals obtained from estimating the corresponding AR-ARCH models.

4. Using the BDS test, Hsieh (1989) found no evidence of in-mean nonlinearities in daily foreign exchange rates after properly specifying their conditional distributions and time variation in conditional volatility. However, using neural networks, Kuan and Liu (1995) provided statistically significant out of sample forecasting improvements over the random walk model for daily exchange rates.
5. The results do not depend on the particular choice of the out-of-sample period, as similar evidence is obtained for different choices of the test set.
6. We also estimated in-sample nonparametric autoregressions up to twelfth order and compared their forecasting performance to those of the linear benchmark models. The in-sample nonparametric fits strictly dominate the linear fits in terms of RMSE predictive accuracy across all autoregressive orders and window sizes considered. The nonparametric fit improves with increasing lag orders and deteriorates with increasing window sizes, thus suggesting overfitting. To preserve space, these results are nor reported here but they are available upon request from the authors.
7. We estimated nonlinear autoregressions up to order twelve. Our evidence is robust as the result remain qualitatively the same for higher autoregression orders. Full results are available upon request from the authors.
8. LeBaron (1992) did provide some forecast improvements for stock returns and foreign exchange rates using the locally unweighted regression method with the level of volatility as the crucial element of conditioning information.

References

Brennan, M., & Schwartz, E. (1979). A continuous time approach to the pricing of bonds, *Journal of Banking and Finance*, 3, 135-155.

Brock, W. A. (1986). Distinguishing random and deterministic systems: An expanded version, *Journal of Economic Theory*, 90, 168-195.

Brock, W., Dechert, W. D., & Scheinkman, J. (1987). A Test for Independence Based on the Correlation Dimension, Economics Working Paper SSRI-8702, University of Wisconsin.

Brock, W., Hsieh, D., & LeBaron, L. (1991). *Nonlinear Dynamics, Chaos and Instability*, Cambridge, MIT Press, MA.

Broze, L., Scaillet, O., & Zakoian, J. (1995). Testing for continuous time models of the short-term interest rate, *Journal of Empirical Finance*, 2, 199-223.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance*, 47, 1209-1227.

Cleveland, W. A. (1979). Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, 74, 829-836.

Cleveland, W. A., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting, *Journal of the American Statistical Association*, 83, 596-610.

Cleveland, W. S., Devlin, S., & Grosse, E. (1988). Regression by local fitting: Methods, properties, and computational algorithms, *Journal of Econometrics*, 37, 87-114.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985a). An intertemporal general equilibrium model of asset prices, *Econometrica*, 53, 363-384.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985b). A theory of term structure of interest rates, *Econometrica*, 53, 385-407.

Dickey, D.A.,. & Fuller, W.A. (1979). Distribution of estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*. 84. 427-431.

Diebold, F. X., & Nason, J. (1990). Nonparametric Exchange Rate Prediction, *Journal of International Economics*, 28, 315-332.

Diebold, F. X., & Rudebusch, G. D. (1991). On the power of Dickey-Fuller tests against fractional alternatives, *Economics Letters*, 35, 155-160.

Dothan, U. L. (1978). On the term structure of interest rates, *Journal of Financial Economics*, 6, 59-69.

Fuller, W. (1976). *Introduction to Statistical Time Series*. New York: John Wiley & Sons.

Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models , *Journal of Time Series Analysis*, 4, 221-238.

Granger, C. W., & Newbold, P. (1986). *Forecasting Economic Time Series* (2nd edition). Academic Press, London and San Diego.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica*, 60, 77-105.

Hodrick, P. J. (1987). Risk, Uncertainty, and Exchange Rates, NBER working paper no. 2429.

Hsieh, D. (1989). Testing for nonlinear dependence in daily foreign exchange rates, *Journal of Business*, 62, 329-368.

Hsieh, D. (1991). Chaos and nonlinear dynamics: Application to financial markets, *Journal of Finance*, 5, 1839-1877.

Hsieh, D., & LeBaron, B. (1988). Finite Sample Properties of the BDS Statistic, Working Paper, University of Chicago and University of Wisconsin at Madison.

Kendall, M., & Stuart, A. S. (1977). *The Advanced Theory of Statistics*, New York: Macmillan.

Kuan, C., & Liu, T. (1995). Forecasting exchange rates using feedforward and recurrent neural networks, *Journal of Applied Econometrics*, 7, S137-S149.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?, *Journal of Econometrics*, 54, 159-178.

Langetieg, T. (1980). A multivariate model of the term structure, *Journal of Finance*, 35, 71-97.

LeBaron, B. (1988). The Changing Structure of Stock Returns, Working Paper, University of Wisconsin.

LeBaron, B. (1992). Forecast improvements using a volatility index, *Journal of Applied Econometrics*, 7, S137-S149.

Longstaff, F.A., & Schwartz, E.S. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model, *Journal of Finance*, 47, 1259-1282.

Meese, R. A., & Rose, A. K. (1990). Nonlinear, nonparametric, nonessential exchange rate estimation, American Economic *Review Papers and Proceedings*, 14, 192-196.

Meese, R. A., & Rose, A. K. (1991). An empirical assessment of nonlinearities in models of exchange rate determination, *Review of Economic Studies*, 80, 603-619.

Merton, M. C. (1973). The theory of rational oprion pricing, *Bell Journal of Economics and Management Science*, 4, 141-183.

Mizrach, B. (1992), Multivariate nearest-neighbor forecasts of EMS exchange rates, *Journal of Applied Econometrics*, 7, S151-S163.

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

Phillips, P. C. B. (1987). Time series regression with a unit root, *Econometrica*, 55, 277-301.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, 75, 335-346.

Richard, S. F. (1978). An arbitrage model of the term structure of interest rates, *Journal of Financial Economics*, 6, 33-57.

Schaefer, S. M., & Schwartz, E. S. (1984). A two-factor model of the term structure: An approximate analytical solution, *Journal of Financial and Quantitative analysis*, 19, 413-424.

Stone, C. J. (1977). Consistent non-parametric regression, *Annals of Statistics*, 5, 595-645.

Vasicek, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics*, 5, 177-188.

Table 1
Summary Statistics of Yield Changes, 1957:02-1988:12

Statistic	Federal Funds	3-Month	6-Month	12-Month	5-Year	10-Year
Mean	0.0160	0.0129	0.0142	0.0117	0.0100	0.0102
Median	0.0300	0.0200	0.0300	0.0300	0.0200	0.0200
Standard Deviation	0.6968	0.5642	0.5406	0.5476	0.3837	0.3210
Skewness	-1.8821***	-1.5677***	-1.5776***	-1.3504***	0.5535***	-0.5728***
Kurtosis	26.6074***	17.0442***	13.9526***	11.2196***	6.1584***	6.2467***
Minimum	-6.6300	-4.6200	-4.2300	-3.9100	-2.0300	-1.7600
Maximum	3.0600	2.6100	2.1700	1.9000	1.8600	1.6100

For the 6-month yield series the sample period is 1959:02 to 1988:12. Kurtosis refers to excess kurtosis. The superscript *** indicates statistical significance at the 1 per cent level.

Table 2
Phillips-Perron Unit Root Tests on Yield Series, 1957:02-1988:12

	Test Statistic						
Yield Series	$Z(a^*)$	$Z(t_{a^*})$	$Z(\phi_1)$	$Z(\tilde{a})$	$Z(t_{\tilde{a}})$	$Z(\phi_2)$	$Z(\phi_3)$
Levels							
Federal Funds	-10.54	-2.34	2.81	-19.89	-3.04	3.15	4.72
3-Month	-9.49	-2.22	2.54	-17.87	-2.96	2.99	4.48
6-Month	-9.15	-2.22	2.54	-15.10	-2.70	2.52	3.79
12-Month	-8.28	-2.10	2.28	-15.78	-2.74	2.58	3.88
5-Year	-4.97	-1.71	1.59	-9.96	-2.04	1.60	2.41
10-Year	-3.79	-1.55	1.39	-7.21	-1.64	1.22	1.83
Differences							
Federal Funds	-222.4	-13.78	86.73	-211.9	-13.16	57.64	86.79
3-Month	-229.2	-13.66	93.22	-229.2	-13.64	61.94	93.24
6-Month	-211.9	-13.78	86.73	-211.9	-13.16	57.64	86.79
12-Month	-215.3	-13.22	87.34	-215.3	-13.20	58.04	87.38
5-Year	-215.7	-13.20	87.04	-215.8	-13.19	57.93	87.21
10-Year	-232.2	-13.77	94.78	-232.4	-13.77	63.18	95.09

For the 6-month yield series the sample period is 1959:02 to 1988:12. The seven different statistics all test for a unit root in the univariate time-series representation for each of the series against a stationary or trend-stationary alternative. In constructing the test statistics the order of serial correlation allowed is 4. Alternative lag structures were also allowed but inference is robust to the order of serial correlation. We used the lag window suggested by Newey and West (1987) to ensure positive semidefiniteness. For more details on the tests see Phillips (1987), and Phillips and Perron (1988). The critical values are as follows (Fuller (1976), and Dickey and Fuller (1979)) with rejections of the null hypothesis indicated by large absolute values of the statistics:

Critical Values	$Z(a^*)$	$Z(t_{a^*})$	$Z(\phi_1)$	$Z(\tilde{a})$	$Z(t_{\tilde{a}})$	$Z(\phi_2)$	$Z(\phi_3)$
10%	-11.3	-2.57	3.78	-18.3	-3.12	4.03	5.34
5%	-14.1	-2.86	4.59	-21.8	-3.41	4.68	6.25
2.5%	-16.9	-3.12	5.38	-25.1	-3.66	5.31	7.16
1%	-20.7	-3.43	6.43	-29.5	-3.96	6.09	8.27

Table 3
Empirical Estimates for the Fractional-Differencing Parameter $\tilde{d}$, 1957:02-1988:12

Yield Series	$\tilde{d}(0.50)$	$\tilde{d}(0.55)$	$\tilde{d}(0.60)$
Federal Funds	0.086 (0.513)	-0.036 (-0.276)	-0.154 (-1.503)‡
3-Month	0.018 (0.105)	0.017 (0.122)	-0.145 (-1.271)
6-Month	-0.031 (-0.179)	0.009 (0.058)	-0.165 (-1.313)‡
12-Month	-0.002 (-0.023)	0.006 (0.069)	-0.108 (-1.153)
5-Year	0.030 (0.259)	0.072 (0.754)	-0.003 (-0.042)
10-Year	0.037 (0.361)	0.170 (1.717)*,##	0.101 (1.190)

Notes: For the 6-month yield series the sample period is 1959:02 to 1988:12. d is the fractional differencing parameter corresponding to the yield series whereas $\tilde{d}$ is the fractional differencing parameter corresponding to the yield-change series $\left(d = 1 + \tilde{d}\right)$ $\tilde{d}(0.50)$, $\tilde{d}(0.55)$ and $\tilde{d}(0.60)$ give the estimates corresponding to the GPH spectral regression of sample size $\nu = T^{0.50}$, $\nu = T^{0.55}$ and $\nu = T^{0.60}$, respectively. The t-statistics are given in parentheses and are constructed imposing the known theoretical error variance of $\pi^2/6$. The superscripts ***, **, * indicate statistical significance for the null hypothesis $\tilde{d} = 0\,(d = 1)$ against the alternative $\tilde{d} \neq 0\,(d \neq 1)$ at the 1, 5, and 10 per cent levels, respectively. The superscripts ###, ##, # indicate statistical significance for the null hypothesis $\tilde{d} = 0\,(d = 1)$ against the one-sided alternative $\tilde{d} > 0\,(d > 1)$ at the 1, 5, and 10 per cent levels, respectively. The superscripts ‡‡‡, ‡‡, ‡ indicate statistical significance for the null hypothesis $\tilde{d} = 0\,(d = 1)$ against the one-sided alternative $\tilde{d} < 0\,(d < 1)$ at the 1, 5, and 10 per cent levels, respectively.

Table 4
BDS Test Results on Yield Changes, 1957:02-1988:12.

	m							
	2		3		4		5	
	ε/σ		ε/σ		ε/σ		ε/σ	
Yield Series	0.5	1.0	0.5	1.0	0.5	1.0	0.5	1.0
Differences								
Federal Funds	10.92	9.91	13.01	11.21	16.02	12.15	19.91	12.81
3-Month	9.96	10.47	12.46	11.35	16.06	12.24	20.24	13.13
6-Month	9.39	10.63	12.24	11.82	16.16	13.19	21.49	14.69
12-Month	10.31	9.91	14.81	12.31	19.92	13.83	28.21	15.69
5-Year	11.31	9.20	16.20	12.32	21.16	14.26	28.78	16.10
10-Year	9.24	8.54	12.92	11.37	17.17	13.41	24.54	15.66
AR-Filtered Differences								
Federal Funds	8.12	7.96	10.63	9.88	13.92	11.70	17.68	13.01
3-Month	7.32	7.66	12.09	9.82	16.62	11.20	23.42	13.16
6-Month	6.58	7.56	7.64	8.19	9.36	9.14	12.47	10.70
12-Month	9.16	9.23	10.99	10.07	15.86	11.78	24.72	13.95
5-Year	7.80	5.96	13.17	8.65	20.23	10.60	30.51	12.24
10-Year	6.17	5.66	10.22	7.99	14.73	9.56	21.16	11.26
AR-ARCH-Filtered Differences								
Federal Funds	1.45	0.38	0.09	-0.45	-0.03	-0.85	0.63	-0.98
3-Month	0.51	0.22	0.32	-0.14	0.26	-0.34	0.02	-0.60
6-Month	1.22	0.20	1.16	0.22	2.54	0.63	3.69	1.19
12-Month	-0.40	-0.76	-1.48	-1.49	-1.28	-1.31	-1.66	-0.95
5-Year	-1.11	-1.75	-0.95	-1.85	0.92	-1.02	4.67	-0.35
10-Year	-2.28	-1.70	-2.42	-2.00	-2.54	-2.27	-3.07	-2.37
5% Critical Values for Linear Series	-2.64 2.98	-2.15 2.27	-2.92 3.23	-2.17 2.37	-3.37 3.84	-2.17 2.39	-4.11 4.98	-2.18 2.56
5% Critical Values for ARCH Residuals	-2.00 1.95	-1.75 1.51	-1.95 2.08	-1.52 1.39	-2.19 2.46	-1.35 1.33	-2.74 3.25	-1.29 1.46

For the 6-month yield series the sample period is 1959:02 to 1988:12. The BDS (m, ε) tests for i.i.d. where m is the embedding dimension and is distance, set in terms of the standard deviation of the data(σ) to 0.5 and 1.0 standard deviations. The AR-filtered series are the residual series obtained from fitting an AR model to each of the yield-change series. The lag order is chosen according to AIC and is 13, 20, 19, 19, 6, and 22 for the Federal Funds, 3-month, 6-month, 12-month, 5-year and 10-year yield series, respectively. The ARCH orders chosen are 3, 4, 2, 2, 2, and 4 for the Federal Funds, 3-month, 6-month, 12-month, 5-year, and

10-year yield-change series, respectively. Finite-sample critical values for the BDS test applied to linear series (yield differences and AR-filtered yield differences) are the 95 per cent quantiles reported by Brock et al. for 250 observations (1991, Table C2). Finite-sample critical values for the BDS test applied to AR-ARCH standardized residuals are approximated by the 95 per cent quantiles reported by Brock et al. on GARCH (1,1) standardized residuals for 500 observations (1991, Table F2).

Table 5
Out-of-sample RMSE from Alternative Models for Predicting Federal Funds Rate Changes, 1989:01-1993:12

Window Size	Lags=1	Lags=2	Lags=3	Lags=4	Lags=5	Lags=6
0.10	0.1869	0.2050	0.1841	0.1912	0.2130	0.2189
0.20	0.1831	0.1908	0.1810	0.1826	0.2011	0.1966
0.30	0.1861	0.1897	0.1847	0.1785	0.1954	0.2004
0.40	0.1824	0.1898	0.1823	0.1827	0.1926	0.1885
0.50	0.1807	0.1935	0.1786	0.1770	0.1876	0.1857
0.60	0.1799	0.1868	0.1775	0.1793	0.1871	0.1886
0.70	0.1748	0.1837	0.1848	0.1871	0.1857	0.1826
0.80	0.1775	0.1910	0.1790	0.1801	0.1813	0.1874
0.90	0.1824	0.1889	0.1795	0.1752	0.1717	0.1758
AR(13)	0.2033					
RW	0.2216					

Window size refers to the fraction of in-sample observations which are chosen as nearest neighbors in the locally weighted regression (LWR). Lags stands for the lag order in the nonlinear autoregression estimated by the LWR method. AR(k) stands for a linear autoregression of order k, which is chosen by the AIC criterion. RW stands for random walk with drift. RMSE denotes root mean square error. Those RMSEs obtained from the LWR method which are lower than the ones obtained from the AR model are underlined.

Table 6
Out-of-sample RMSE from Alternative Models for Predicting 3-Month T-bill Yield Changes, 1989:01-1993:12

Window Size	Lags=1	Lags=2	Lags=3	Lags=4	Lags=5	Lags=6
0.10	0.1961	0.1897	0.1861	0.1747	0.1753	0.1799
0.20	0.1931	0.1864	0.1852	0.1712	0.1820	0.1799
0.30	0.1889	0.1858	0.1690	0.1830	0.1740	0.1834
0.40	0.1848	0.1833	0.1676	0.1785	0.1757	0.1742
0.50	0.1869	0.1811	0.1737	0.1846	0.1715	0.1754
0.60	0.1853	0.1807	0.1730	0.1782	0.1737	0.1772
0.70	0.1827	0.1787	0.1752	0.1812	0.1746	0.1799
0.80	0.1806	0.1821	0.1757	0.1772	0.1780	0.1845
0.90	0.1790	0.1772	0.1782	0.1744	0.1771	0.1892
AR(20)	0.2104					
RW	0.2068					

See notes in Table 5.

Table 7
Out-of-sample RMSE from Alternative Models for Predicting 6-Month T-bill Yield Changes, 1989:01-1993:12

Window Size	Lags=1	Lags=2	Lags=3	Lags=4	Lags=5	Lags=6
0.10	0.1998	0.1893	0.2294	0.2222	0.2173	0.2219
0.20	0.1923	0.1911	0.2046	0.2000	0.2039	0.2221
0.30	0.1861	0.1882	0.2019	0.2030	0.1968	0.2109
0.40	0.1831	0.1945	0.1965	0.1958	0.1867	0.1942
0.50	0.1822	0.1903	0.1877	0.1873	0.1804	0.1881
0.60	0.1830	0.1868	0.1874	0.1866	0.1861	0.1862
0.70	0.1827	0.1879	0.1879	0.1897	0.1946	0.1905
0.80	0.1818	0.1886	0.1870	0.1864	0.1839	0.1901
0.90	0.1879	0.1837	0.1842	0.1831	0.1825	0.1965
AR(19)	0.2104					
RW	0.2109					

See notes in Table 5.

Table 8
Out-of-sample RMSE from Alternative Models for Predicting 12-Month T-bill Yield Changes, 1989:01-1993:12

Window Size	Lags=1	Lags=2	Lags=3	Lags=4	Lags=5	Lags=6
0.10	0.2111	0.2082	0.2195	0.2262	0.2334	0.2626
0.20	0.2200	0.2166	0.2252	0.2185	0.2195	0.2370
0.30	0.2166	0.2191	0.2166	0.2280	0.2195	0.2298
0.40	0.2176	0.2193	0.2165	0.2250	0.2266	0.2168
0.50	0.2165	0.2208	0.2164	0.2219	0.2091	0.2123
0.60	0.2158	0.2176	0.2153	0.2140	0.2105	0.2126
0.70	0.2163	0.2198	0.2182	0.2166	0.2025	0.2177
0.80	0.2144	0.2206	0.2187	0.2180	0.2155	0.2203
0.90	0.2202	0.2181	0.2158	0.2218	0.2163	0.2282
AR(19)	0.2568					
RW	0.2427					

See notes in Table 5.

Table 9
Out-of-sample RMSE from Alternative Models for Predicting 5-Year Bond Yield Changes, 1989:01-1993:12

Window Size	Lags=1	Lags=2	Lags=3	Lags=4	Lags=5	Lags=6
0.10	0.2275	0.2401	0.2516	0.2606	0.2621	0.2669
0.20	0.2310	0.2326	0.2461	0.2496	0.2490	0.2430
0.30	0.2263	0.2300	0.2376	0.2390	0.2412	0.2404
0.40	0.2270	0.2258	0.2359	0.2310	0.2321	0.2326
0.50	0.2253	0.2231	0.2336	0.2319	0.2272	0.2226
0.60	0.2282	0.2223	0.2236	0.2279	0.2245	0.2211
0.70	0.2274	0.2197	0.2268	0.2263	0.2249	0.2268
0.80	0.2287	0.2206	0.2259	0.2242	0.2274	0.2375
0.90	0.2287	0.2209	0.2234	0.2246	0.2307	0.2356
AR(6)	0.2326					
RW	0.2562					

See notes in Table 5.

Table 10
Out-of-sample RMSE from Alternative Models for Predicting 10-Year Bond Yield Changes, 1989:01-1993:12

Window Size	Lags=1	Lags=2	Lags=3	Lags=4	Lags=5	Lags=6
0.10	0.1958	0.2108	0.2036	0.2129	0.2044	0.2145
0.20	0.1964	0.2022	0.2005	0.2114	0.2019	0.2110
0.30	0.1994	0.2041	0.2089	0.1982	0.1961	0.2020
0.40	0.1983	0.1988	0.2081	0.1980	0.2007	0.2007
0.50	0.2023	0.1985	0.1965	0.1995	0.1972	0.1938
0.60	0.1996	0.1931	0.1993	0.1982	0.1972	0.1934
0.70	0.1998	0.1952	0.1979	0.1993	0.1963	0.1960
0.80	0.1997	0.1953	0.1959	0.1983	0.1998	0.2008
0.90	0.2004	0.1919	0.1957	0.1950	0.2000	0.1988
AR(22)	0.2068					
RW	0.2195					

See notes in Table 5.

Table 11
Granger-Newbold Test Results for Out-of-sample Forecasting Performance

Yield Series and Window Size	Lags=1	Lags=2	Lags=3	Lags=4	Lags=5	Lags=6
Federal Funds						
0.10	-0.393	0.827	-0.380	0.185	0.647	0.917
0.20	-0.194	-0.026	-0.416	-0.674	0.047	0.350
0.30	0.110	-0.096	0.239	-0.668	0.012	0.727
0.40	-0.114	-0.008	0.119	-0.045	0.094	-0.491
0.50	-0.165	0.310	-0.067	-0.771	-0.140	-0.548
0.60	-0.406	0.014	-0.164	-0.873	-0.236	-0.468
0.70	-0.914	-0.149	0.157	-0.343	-0.240	-0.782
0.80	-0.450	0.007	-0.483	-0.485	-0.475	-0.402
0.90	-0.129	0.300	-0.089	-0.827	-0.966	-1.094
3-Month Bill						
0.10	-1.043	-0.825	-1.034	-1.849**	-1.261	-1.338*
0.20	-1.385*	-1.624*	1.331*	-2.206**	-1.402*	-2.431***
0.30	-1.724**	-1.871**	-2.857***	-1.742**	-2.352***	-2.235**
0.40	-1.973**	-2.018**	-2.854***	-2.023**	-2.160**	-2.780***
0.50	-1.780**	-1.890**	-2.485***	-1.386*	-2.338***	-2.334***
0.60	-1.958**	-1.769**	-2.228**	-1.674**	-2.158**	-2.449***
0.70	-1.990**	-1.854**	-2.172**	-1.310*	-2.339***	-2.553***
0.80	-1.882**	-1.470*	-2.088**	-1.717**	-2.077**	-2.771***
0.90	-2.286**	-2.323**	-2.093**	-1.696**	-1.310*	-1.519*
6-Month Bill						
0.10	-0.271	-1.876**	0.549	0.920	0.552	0.583
0.20	-0.828	-1.494*	-0.863	-0.496	0.770	-0.123
0.30	-1.333*	-1.800**	-1.013	-0.446	-1.298*	-1.086
12-Month Bill						
0.10	-1.768**	-2.567***	-1.545*	-0.964	-0.697	0.270
0.20	-1.421*	-2.071**	-1.542*	1.503*	-1.343*	-0.980
0.30	-1.627*	-2.022**	-2.167**	-1.058	-1.440*	-1.234
0.40	-1.584*	-1.925**	-2.143**	-1.245	-0.977	-1.879**
0.50	-1.612*	-1.997**	-2.067**	-1.337*	-1.841**	-2.151**
0.60	-1.685**	-2.068**	-2.080**	-1.775**	-1.855**	-2.238**
0.70	-1.669**	-1.791**	-1.757**	-1.593*	-1.709**	-1.912**
0.80	-1.794**	-1.784**	-1.763**	-1.652**	-1.633*	-1.716**
0.90	-1.581*	-1.956**	-1.713**	-1.569*	-1.765**	-1.472*
5-Year Bond						
0.10	-0.004	0.917	1.770	2.505	1.861	1.722
0.20	0.191	0.232	1.358	1.649	1.367	0.996
0.30	-0.241	0.028	0.872	0.886	1.170	1.024
0.40	-0.181	-0.299	0.767	0.362	0.616	0.598
0.50	-0.264	-0.571	0.617	0.472	0.216	-0.303
0.60	-0.026	-0.653	-0.359	0.105	-0.050	-0.468

(sambungan Jadual 11)

0.70	-0.131	-0.920	-0.035	-0.101	-0.006	0.195
0.80	0.013	-1.034	-0.067	-0.312	0.158	1.415
0.90	0.019	-1.039	-0.439	-0.390	0.338	1.213
10-Year Bond						
0.10	-1.461*	-0.638	-1.074	-0.301	-0.774	-0.194
0.20	-1.426*	-1.241	-1.386*	-0.367	-0.931	-0.355
0.30	-1.210	-1.074	-0.554	-1.303*	-1.325*	-1.079
0.40	-1.366*	-1.526*	-0.646	-1.305*	-0.981	-0.995
0.50	-1.048	-1.629*	-1.571*	-1.197	-1.293*	-1.475*
0.60	-1.218	-2.058**	-1.223	-1.297*	-1.289*	-1.670**
0.70	-1.210	-1.842**	-1.346*	-1.363*	-1.448*	-1.541*
0.80	-1.232	-1.833**	-1.476*	-1.379*	-1.221	-1.141
0.90	-1.117	-2.174**	-1.577*	-1.638*	-1.307*	-1.563*

The Granger-Newbold test (Granger and Newbold, 1986) is based on the correlation coefficient of the sums and differences of the forecast errors. We test the null hypothesis $H_0: r = 0$ against the one-sided alternative that $r < 0$, that is, that the LWR model has superior forecasting performance over the AR model. The null hypothesis is based on the approximation that $w = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \sim N\left(0, \left(\frac{1}{T-3}\right)\right)$ where is the number of forecasts, 48 in our case.