NORTH-HOLLAND

International Review of Economics and Finance
10 (2001) 353–368

International Review of Economics & Finance

Nonlinear predictability of stock market returns: Evidence from nonparametric and threshold models

David G. McMillan*

Department of Economics, University of St. Andrews, St. Andrews, Fife KY16 9AL, UK

Received 20 September 2000; received in revised form 27 February 2001; accepted 15 May 2001

Abstract

Recent empirical evidence suggests that stock market returns are predictable from a variety of financial and macroeconomic variables. However, with two exceptions this predictability is based upon a linear functional form. This paper extends this research by considering whether a nonlinear relationship exists between stock market returns and these conditioning variables, and whether this nonlinearity can be exploited for forecast improvements. General nonlinearities are examined using a nonparametric regression technique, which suggest possible threshold behaviour. This leads to estimation of a smooth-transition threshold type model, with the results indicating an improved in-sample performance and marginally superior out-of-sample forecast results. © 2001 Elsevier Science Inc. All rights reserved.

JEL classification: G12; G13

Keywords: Stock market returns; Nonparametric regression; STARX model; Predictability

1. Introduction

An increasing amount of empirical evidence points to the conclusion that stock market returns can be predicted by a range of financial and macroeconomic variables. Recent studies (e.g. Balvers, Cosimano, & McDonald, 1990; Breen, Glosten, & Jagannathan, 1990;

* Tel.: +44-1334-462420; fax: +44-1334-462444.
E-mail address: dgm6@st-andrews.ac.uk (D.G. McMillan).

PII: S1059-0560(01)00093-4

Campbell, 1987; Campbell & Hamao, 1992; Cochrane, 1991; Fama & French, 1989; Ferson & Harvey, 1993; French, Schwert, & Stambaugh, 1987; Glosten, Jagannathan, & Runkle, 1993; Keim & Stambaugh, 1986; Pesaran & Timmerman, 1995) have shown that this conclusion holds across a variety of stock markets and time horizons despite its implication for market efficiency.[1]

In a recent advancement Pesaran and Timmerman (1995, 2000) using a linear recursive modelling strategy examined the robustness of predictability of US and UK stock market returns by simulating the behaviour of investors who search in 'real time' for a model that can forecast stock returns. In each time period the regression model is reviewed, with some variables always included in the regression model, thus being viewed a priori important while others are selected according to certain criteria such as the Akaike and Schwarz information criteria. These variables are then used to perform one-step ahead forecasts. Thus, an investor is open-minded as to which variables should be included in the information set. Pesaran and Timmermann reported evidence of predictability in both US and UK stock markets which could have been exploited by investors.

However, whilst in each time period the regression variables are reviewed for inclusion in the forecast model, the functional form of the regression is not reexamined. Indeed the vast majority of extant work examines the predictability of stock returns by financial and macroeconomic variables using a linear regression framework. This despite increasing evidence of nonlinear behaviour in asset returns (e.g., Abhyankar, Copeland, & Wong, 1997, provide a summary of recent evidence of nonlinearity). Recent exceptions to this is the work of Qi (1999) who uses a neural network method, and thus provides flexibility in the choice between linear and nonlinear models, to examine the predictability of US stock returns, and Perez-Quiros and Timmermann (2000) who use a Markov switching model to examine returns in large and small US firms. Thus, while in the Pesaran and Timmermann framework the investor knows the functional form but is open-minded to the conditioning set, an alternative is for the investor to select the set of financial and macroeconomic variables, whilst being open-minded to the underlying specification. This paper continues the latter theme and following estimation of a linear model, examines the relationship between stock market returns and various financial and macroeconomic variables using a model-free nonparametric estimator. Following this procedure we then examine the resulting plots to see if this suggests any evidence of nonlinear form, and from this we tentatively propose a parametric nonlinear model of smooth transition threshold form. More specifically, a model based upon the smooth transition regression (STR) and autoregressive STR (STAR) type models is estimated (see Chan & Tong, 1986; Granger & Teräsvirta, 1993; Teräsvirta, 1994; Teräsvirta & Anderson, 1992), referred to here as a STARX model as exogenous variables are used as explanatory variables, but with an autoregressive transition variable.

[1] Poterba and Summers (1988) argue that the idea that predictable components in share prices arise as a result of rational variation in expected excess returns is not inconsistent with the concept of market efficiency. While Balvers et al. (1990) construct a general equilibrium model relating asset returns to macroeconomic fluctuations in a context that is consistent with efficient markets.

The remainder of this paper is organised as follows, Section 2 briefly describes the data, and examines in greater detail the empirical methodology, including the nonparametric and nonlinear regression models. Section 3 presents our results and Section 4 summarises and concludes.

2. Data and empirical methodology

2.1. Data

The stock market index data analysed here is S&P 500 monthly index returns from the period January 1970 to March 1995, with the sample period from April 1995 to March 2000 being used in the forecasting exercise.[2] While the following monthly financial and macroeconomic data are used to attempt to predict the returns process, the 3-month Treasury bill (T-bill), the 12-month T-bill, unemployment, industrial production, consumer price index and money supply M1. All data were tested for the presence of unit roots using the test Augmented Dickey–Fuller tests. The results suggest a single unit root in each series, except the stock market index returns series, where a unit root is present in the levels (price) data, thus to ensure stationarity all relevant data are differenced.[3]

2.2. Linear model

We begin the examination of the data by using a standard linear regression model of the following form:

$$x_t = a_0 + \sum_{i=1}^{m} a_i z_{t-i} + \varepsilon_t \tag{1}$$

where x_t is stock market returns and z_t is a vector which contains the exogenous variables described in the previous section. Initially we consider a lag length of six for each of the right-hand side variables, with lags being eliminated on the basis of individual significance test, while joint significance tests are performed on the final specific model, information criteria such as the AIC and BIC were also used to inform appropriate lag lengths decisions.

2.3. Nonparametric estimation

To consider whether linearity is the appropriate functional form we proceeded to consider nonparametric regression of the returns processes against the significant exogenous variables

[2] Analysis was also conducted on Dow Jones Industrial Average (DJIA) monthly returns, with the results being qualitatively similar to those reported for the S&P 500 returns and are noted in subsequent footnotes.

[3] Unit root test results are suppressed for space consideration, but available upon request from the author.

as identified using the linear model above.[4] The nonparametric procedure is based upon the estimation of a probability density function, pioneered by Rosenblatt (1956), with applications to regression work by Nadaraya (1964), Stone (1977) and Watson (1964).[5] If we again consider our conditional mean equation we have Eq. (2):

$$x_t = f(z_{t-i}) + \varepsilon_t \tag{2}$$

where $f(z_{t-i})$ includes the significant variables identified from Eq. (1), although in principle it could include autoregressive and moving average terms. Whilst parametric estimation of x_t involves specifying a specific functional form for $f(\cdot)$, an alternative approach is to estimate the function via some smoothing operation with no functional form specified. The method of estimation chosen here is to use a weighted average as such (Eq. (3)):

$$\hat{x}_t = \sum_{j=1}^{T} w_{jt} z_{t-i}; \quad \sum_{j=1}^{T} w_{jt} = 1 \tag{3}$$

where the weights accorded depend upon the proximity of the points x_t to given z_{t-i} values. Whilst a variety of weighting schemes are available, the scheme chosen here is one of the more popular methods and that largely used in other studies (see Pagan & Schwert, 1990), namely the Nadaraya–Watson estimator (Eq. (4)):[6]

$$f(x) = \sum_{i=1}^{n} x_i K(z_i - z/h) / \sum_{i=1}^{n} K(z_i - z/h) \tag{4}$$

where $K(\cdot)$ is the kernel weighting function, and h defines the bandwidth or 'smoothing parameter' which determines the degree of smoothness imposed upon the estimation, and is a function of the sample size ($h \to 0$, $T \to \infty$). Commonly, the kernel is a probability density function such that $K(\cdot) \geq 0$ and $\int K(x)\mathrm{d}x = 1$, while the optimum kernel function and bandwidth selections minimise the integrated mean square error (IMSE). The choice of kernel used here is the Epanechnikov (1969) kernel which is the optimal kernel based

[4] Although in principle variables found to be insignificant under linear estimation could be significant under nonlinear estimation, we follow the reasoning in Granger and Teräsvirta (1993) that if the data generating process is truly nonlinear then fitting a linear model would overfit the data, resulting in more significant parameters than required by the correct nonlinear specification.

[5] Recent books and review articles on nonparametric regression include Delgado and Robinson (1992), Härdle (1990), Prakasa Rao (1983), Silverman (1986) and Ullah (1988).

[6] The earliest was that of Rosenblatt (1956) who introduced the general class of kernel estimators. The kernel estimator is a sum of curves placed at the data points, where the kernel determines the shape of the curves, and the bandwidth is essentially a generalisation of a histogram bandwidth. Other schemes include the basic histogram, where the data are partitioned before estimation, however discontinuities in the histogram prevent estimation of derivatives; and the nearest-neighbour method which ignores the influence of more distant points.

upon a calculus of variations solution to minimising the IMSE of the kernel estimator. The Epanechnikov kernel is given by:

$$K(x_t) = \left(\frac{3}{4\sqrt{5}}\right)\left(1 - \frac{1}{5}x_j^2\right) \qquad \text{if } x_j^2 < 5.0;\ 0 \text{ otherwise} \tag{5}$$

where the general asymptotically unbiased and mean squares consistency of the kernel has been established by Prakasa Rao (1983) for the case of independent observations, and by Robinson (1983) for dependent observations. More specifically, the kernel estimator is consistent under the following conditions (Eq. (6); where all integrals are defined over the range $\{-\infty, \infty\}$):

$$\int K(x)\mathrm{d}x = 1; \quad \int xK(x)\mathrm{d}x = 0; \quad \int x^2K(x)\mathrm{d}x < \infty; \quad \lim_{n\to\infty} h \to 0;$$

$$\lim_{n\to\infty} nh_j \to \infty. \tag{6}$$

That is, the kernel function $K(\cdot)$ is a twice differentiable 'Borel-measurable' bounded real-valued function symmetric about the origin, the bandwidth vector, h, approaches zero as the sample size approaches infinity, and the product of the bandwidth and the sample size approaches infinity as n approaches infinity. The optimum bandwidth selection (i.e., the bandwidth that minimises the IMSE) is given by Eq. (7):

$$h_j^{\text{opt}} = c_j \sigma_j n^{-1/p+4} \tag{7}$$

where c_j refers to a constant scaling factor that depends upon the kernel function $K(\cdot)$ and on the underlying data process, σ_j is the standard deviation of x and p is the number of regressors. Previous nonparametric studies have imposed a value for c in accordance with that suggested by Silverman (1986) to approximate the optimal choice of bandwidth. However, the value of c is strictly data dependent and so we use an automatic bandwidth selection procedure that has been shown to minimise the IMSE, namely the leave-one-out cross-validation procedure, which is defined in Eq. (8):[7]

$$\mathrm{CV}(h) = \sum_{i=1}^{n} (y_i - \hat{f}(x_{-i}, c))^2 \tag{8}$$

where $\hat{f}(x_{-i}, c)$ denotes the leave-one-out estimator evaluated for a particular value of c, and follows from Eq. (5) with the i-th observation excluded. It has been shown (Stone, 1974, 1984) that, asymptotically, the bandwidth that minimises the leave-one-out CV function, CV(h), also minimises the IMSE.

[7] The cross-validation method of bandwidth choice relies on the established principle of out-of-sample predictive validation. The basic algorithm involves removing any single value of x_i from the sample and computing the conditional mean at the x_i from the remaining sample values, and choosing h such that the IMSE is at the minimum.

2.4. STARX model

In order to attempt to provide some nonlinear parametric form for examining the predictability of stock market returns using financial and macroeconomic data, we consider a version of the general class of STR and STAR models (see Chan & Tong, 1986; Granger & Teräsvirta, 1993; Teräsvirta, 1994; Teräsvirta & Anderson, 1992) that allows for smooth transition between regimes of behaviour. This model is favoured over the simple threshold models which imposes an abrupt switch in parameter values, first, because only if all traders act simultaneously will this be the observed outcome, for a market of many traders acting at slightly different times a smooth transition model is more appropriate. Second, the STAR model allows different types of market behaviour depending on the nature of the transition function. In particular the logistic function allows differing behaviour depending on whether returns are positive or negative, while the exponential function allows differing behaviour to occur for large and small returns regardless of sign. The former function is motivated by considerations of the general state of the market, while the latter function may be motivated by considerations of market frictions, such as transactions costs, which create a band of price movements around the equilibrium price, with arbitrageurs only actively trading when deviations from equilibrium are sufficiently large.[8] Given that we are attempting to use exogenous variables to explain returns but with an autoregressive transition variable this model is termed STARX (smooth transition threshold autoregressive-exogenous) and the model is given by:

$$x_t = \pi_0 + \sum_{i=1}^{p} \pi_i z_{t-1} + \left(\theta_0 + \sum_{i=1}^{p} \theta_i z_{t-1}\right) F(x_{t-d}) + \varepsilon_t \tag{9}$$

where $F(x_{t-d})$ is the transition function. As already stated, two transition functions are considered. The logistic function is given as follows, with the full model thus referred to as a Logistic STARX (or LSTARX) model:

$$F(x_{t-d}) = (1 + \exp(-\gamma(x_{t-d} - c)))^{-1}; \qquad \gamma > 0 \tag{10}$$

which allows a smooth transition between the differing dynamics of positive and negative returns, where d is the delay parameter, γ the smoothing parameter, and c the transition parameter. This function allows the parameters to change monotonically with x_{t-d}. As $\gamma \to \infty$, $F(x_{t-d})$ becomes a Heaviside function: $F(x_{t-d}) = 0$, $x_{t-d} \leq c$, $F(x_{t-d}) = 1$, $x_{t-d} > c$, and Eq. (9) reduces to a TARX(p) model, As $\gamma \to 0$, Eq. (9) becomes a linear model of order p.

[8] An alternative ESTAR motivation is provided by consideration of market depth, whereby the process by which the market can clear reasonable quantities of stock at market prices may differ from the process required to trade large quantities of stock outside the range of price necessary to clear the market.

The second transition function considered is exponential, with the resulting model referred to as the Exponential STARX (or ESTARX) model:

$$F(x_{t-d}) = 1 - \exp(-\gamma(x_{t-d} - c)^2); \quad \gamma > 0. \tag{11}$$

whereby the parameters in Eq. (11) change symmetrically about c with x_{t-d}. If $\gamma \to \infty$ or $\gamma \to 0$ the ESTARX model becomes linear. This model implies that the dynamics of the middle ground differ from those of the larger returns. The ESTAR model is a generalisation of the regular exponential autoregressive (EAR) model of Haggan and Ozaki (1981), where $\theta_0 = c = 0$, this generalisation making the EAR model location invariant. The ESTARX model, which identifies differing behaviour resulting from larger and small trades, may therefore capture the effects of transactions costs on trader behaviour or market depth. For example, whether deviations from the equilibrium price are sufficiently large to allow profitable trade. Alternatively, whether a larger range of trades can be represented by the same process, in which case the market may be said to be ‘deep’, or whether the market is characterised by limited depth, in which case the middle regime of the ESTARX model may be narrow.

To specify the STARX models we use the variables identified as significant from Eq. (1), while a delay parameter of one is adopted. The rationale for this is that we would expect the stock market to react within 1 month to news that alters regime.[9] Finally, estimation of STAR models, and in particular the smoothing parameter γ, has in practice been problematic (see Granger & Teräsvirta, 1993; Teräsvirta, 1994; Teräsvirta & Anderson,1992). In the LSTAR model, a large value for γ results in a steep slope of the transition function at c, and a large number of observations in the neighbourhood of c are therefore required to estimate γ accurately. A result of this is that convergence of γ may be slow, with relatively large changes in γ having only a minor effect upon the shape of the transition function. A solution to this, suggested by Granger and Teräsvirta (1993), Teräsvirta (1994) and Teräsvirta and Anderson (1992), is to scale the smoothing parameter, γ, by the standard deviation of the transition variable, and similarly in the ESTAR model to scale by the variance of the transition variable. Thus, the LSTARX and ESTARX model becomes, respectively, Eqs. (10′) and (11′):

$$F(x_{t-d}) = (1 + \exp(-\gamma(x_{t-d} - c)/\sigma(x_{t-d})))^{-1} \tag{10′}$$

$$F(x_{t-d}) = (1 - \exp(-\gamma(x_{t-d} - c)^2/\sigma^2(x_{t-d}))). \tag{11′}$$

Estimation of the STARX models is by nonlinear least squares. If convergence is obtained, the validity of the model is then evaluated. This includes examination of the parameter values,

[9] A more formal procedure for specifying STAR models is outlined in Granger and Teräsvirta (1993), Teräsvirta (1994) and Teräsvirta and Anderson (1992). The results of which support both the presence of STAR-type nonlinearity and a delay parameter of one, full results are available upon request.

Table 1
Conditional mean linear model estimates and residual tests

	Linear model (S&P 500)	Residual tests	
a0	.0093* (.0025)		
3-month T-bill (− 4)	.01660* (.0054)	AIC	−3.4543
12-month T-bill (− 1)	− .0105* (.0046)	LM_1	0.30 (0.58)
12-month T-bill (− 4)	− .0277* (.0062)	LM_6	5.76 (0.45)
Unemployment (− 1)	.0257* (.0128)	Het	1.77 (0.99)

For equation specification see Eq. (1). $LM_{1/6}$ refer to the Breusch–Godfrey Serial Correlation LM Test, while Het is the White heteroscedasticity test. All tests statistics are the chi-squared statistic with the associated *P* value in parentheses.

* Denotes 5% significance.

particularly ensuring that c, the transition value, is within the range of $\{x_t\}$, and testing for the significance of the explanatory terms. Additionally, the Akaike and Schwarz criterion can be used to guide the selection of competing models (see the examples in Teräsvirta, 1994).[10]

3. Empirical results

Table 1 presents the empirical results for the linear regression model, Eq. (1). As noted above, a lag length of up to six was initially considered, with restrictions made on the basis of individual and joint significance tests and information criteria. The results show that for S&P 500 returns the fourth lag of the 3-month Treasury bill, the first and fourth lag of the 12-month Treasury bill, and the first lag unemployment have significant predictive power, while there are no significant lags of industrial production, CPI, and MI.[11] Table 1 also provides some simple specification diagnostics (the Akaike information criterion) and residual serial correlation and heteroscedasticity tests, which are insignificant.

The analysis conducted so far presumes a linear structure between returns and the lagged financial and macroeconomic variables. However, there may exist a nonlinear relationship between returns and these significant variables. For this purpose we conduct a series of

[10] Further examination of the models can be conducted through examining the dynamic properties of the model. First, computing the roots of the characteristic polynomials corresponding to $F(x_{t-d})=0$ and $F(x_{t-d})=1$, we examine the dynamic properties of each regime, and second, evaluating the long-run dynamic properties of the model. This latter procedure can only be performed numerically, where data are generated from the model in question after setting the error term equal to zero, with a sequence of observed values of the series acting as starting values. This procedure could result in a unique single point stable equilibrium, a limit cycle where a set of values repeat themselves perpetually, or diverge (in which case the model is rejected). A final dynamic case is that the model generates chaotic realisations, in which case a small change in the initial values results in divergent, though stable, limit points.

[11] For DJIA returns the second and fifth lag of the 3-month Treasury bill, the first, second, fifth and sixth lag of the 12-month Treasury bill, and the first lag CPI are significant, while there are no significant lags of industrial production, unemployment and MI. Thus, according to our results here there is no predictive power in either industrial production or money supply for stock market returns.

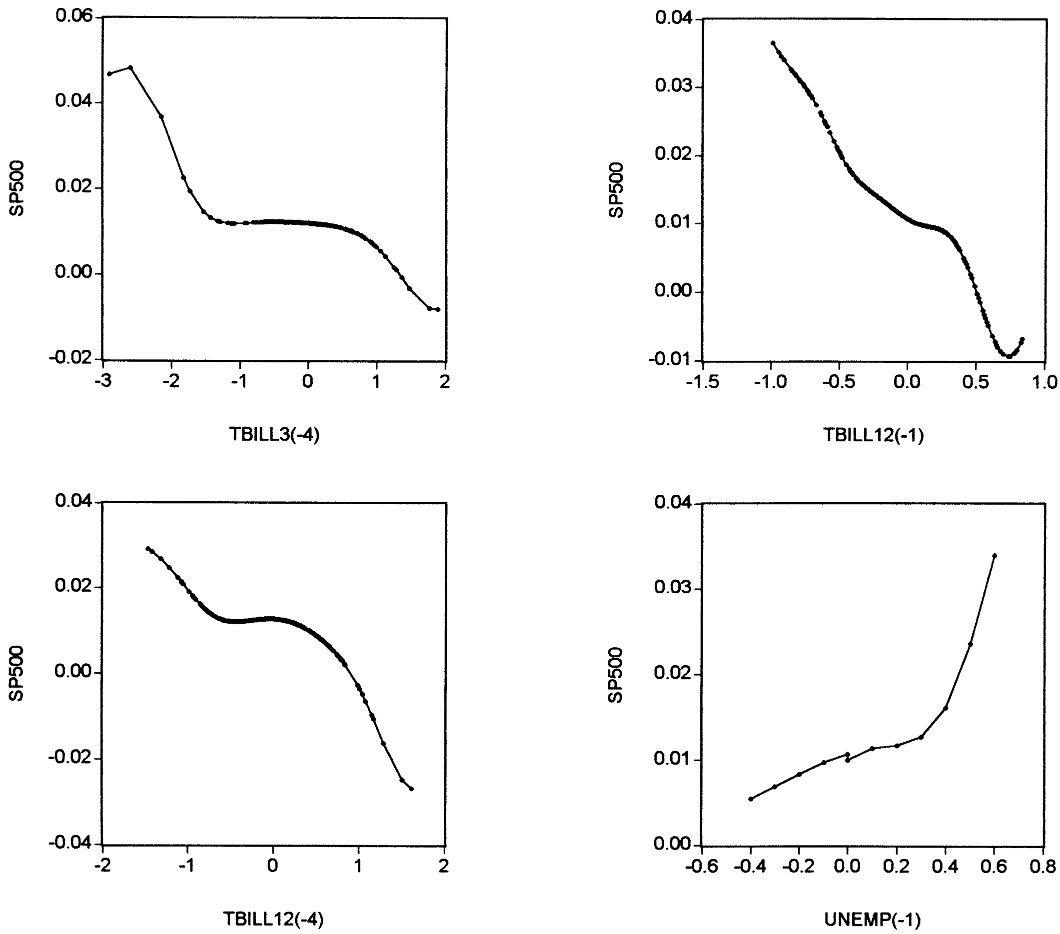

Fig. 1. S&P 500 nonparametric plots.

nonparametric regressions (as described above in Section 2.3) and plot the estimated conditional mean to examine the relationship between stock returns and the identified significant financial and macroeconomic variables. These plots are reported in Fig. 1 and appear to show a nonlinear relationship between the returns on the stock index and returns on the T-bills, except the first lag of the 12-month T-bill. In general, a negative relationship is observed, but with a middle horizontal regime. For unemployment, however, there appears to be a positive linear relationship.[12]

[12] The corresponding plots for the DJIA returns are available upon request from the author. However, examining these plots a similar pattern appears with a negative relationship between the returns and the interest rates variables, but with a middle regime, suggesting some nonlinearity. While for the noninterest rate series, in this case the CPI series, the relationship appears to be a linear one.

From this nonparametric analysis it appears that a nonlinear relationship may indeed exist between stock market returns and financial variables, notably interest rate series. These plots suggest that different regimes of behaviour may exist, with a middle regime that differs from the outer regimes, perhaps indicating some type of threshold effect. With this mind, we tentatively examine a threshold model similar in nature to the STAR model of Chan and Tong (1986), Granger and Teräsvirta (1993), Teräsvirta (1994) and Teräsvirta and Anderson (1992), which we term here STARX as the model contains exogenous variables, with only the delay parameter being an autoregressive term. These estimated models as described in Section 2.4 and Eqs. (9)–(11) are reported in Table 2.

Table 2 shows the nonlinear models for S&P 500 returns. Only the 12-month T-bill series appears in both regimes, whilst the unemployment series appears in the upper and outer regimes of the LSTARX and ESTARX model, respectively, the 3-month T-bill series appears in the lower and middle regime. In both models all lagged variable coefficients are significant at the 5% level, except the fourth lag of the 12-month T-bill series in the LSTARX model which is only significant at the 10% level. The constants are insignificant, except the middle regime constant for the ESTARX model. The speed of transition between regimes parameter, γ, is particularly large for the LSTARX model, suggesting very quick transition between regimes, similar to a standard threshold model, although it is insignificant. The speed parameter for the ESTARX model, although lower, is statistically significant. Finally, the threshold parameter is very similar between both models, suggesting regime shifting at similar returns.[13] Fig. 2 presents graphical evidence of these results, with the transition functions of both models plotted. Evident in these figures is the rapid speed of adjustment between regimes for the LSTARX model, with slower adjustment noticeable for the ESTARX model. Residual tests show no evidence of remaining serial correlation or heteroscedasticity in either model, while the Akaike information criterion supports both nonlinear models over the linear model, and the ESTARX model over the LSTARX model.

In sum, the nonparametric evidence suggested possible nonlinear threshold behaviour between stock market returns and financial variables, in particular interest rates. Estimation of a threshold model variant appears to support this contention, with nonlinear effects particularly noticeable in 12-month T-bill returns, while the threshold parameter is significant for both models.

Given estimation of both linear and nonlinear models it is important to compare the in-sample and out-of-sample performance of these models. Table 3 presents the goodness of fit

[13] The results for the LSTARX model for DJIA returns show that all the exogenous variables appear in the lower regime, except the second lag of the 3-month T-bill, while the same lag, together with the first and sixth lag of the 12-month T-bill and the first lag of the CPI series are insignificant in the upper regime, with all other variables significant. For the ESTARX model all exogenous series, except the second lag of the 3-month T-bill and CPI, appear in the outer regime, while only the CPI series appears in the middle regime, thus the second lag of the 3-month T-bill disappears from the estimated model. As with S&P 500 returns the constant are insignificant, while in contrast, the transition parameter is only significant for the LSTARX model, although again it is larger. Finally, the threshold parameters are significant and of similar magnitude.

Table 2
Conditional mean STARX model estimates

	Nonlinear models		Residual tests LSTAR then ESTAR	
	LSTAR	ESTAR		
π_0	.0233 (.0228)	.0484* (.0215)	LSTARX	
3-month T-bill (− 4)	.0161* (.0054)	.0166* (.0055)	AIC	− 3.4741
12-month T-bill (− 1)	− .0614* (.0158)	− .0859* (.0120)	LM_1	0.76 (0.38)
12-month T-bill (− 4)	− .1132* (.0459)	− .1889* (.0230)	LM_6	7.64 (0.27)
θ_0	− .0134 (.0229)	− .0389 (.0219)	Het	13.67 (0.19)
12-month T-bill (− 1)	.0549* (.0164)	.0800* (.0128)	ESTARX	
12-month T-bill (− 4)	.0876** (.0464)	.1648* (.0236)	AIC	− 3.4931
Unemployment (− 1)	.0296* (.0126)	.0287* (.0128)	LM_1	0.00 (0.97)
γ	39.614 (199.8947)	1.8851* (0.8058)	LM_6	5.88 (0.44)
c	− .0726* (.0163)	− .1079* (.0056)	Het	6.50 (0.77)

For equation specification see Eqs. (9)–(11). $LM_{1/6}$ refer to the Breusch–Godfrey Serial Correlation LM Test, while Het is the White heteroscedasticity test. All tests statistics are the chi-squared statistic with the associated *P* value in parentheses.

* Denotes 5% significance.

** Denotes 10% significance.

measures using the standard methods of root mean squared error, mean absolute error and mean absolute percentage error. These statistics were computed using a series of recursive estimates and one-step ahead forecasts, such that the estimation period rolls forward each period so that the information used in each forecast would be the same information available to the investor. The results suggest that these nonlinear models provide a better fit to the data in-sample, with all the forecast statistics being lower for the two nonlinear models than the linear model for S&P 500 returns. Of the two nonlinear models, the ESTARX model outperforms the LSTARX model. These results indicate that the nonlinear models considered are able to account for the substantial nonlinearity inherent in the returns series better than the linear model.

The second part of Table 3 presents the same exercise for the out-of-sample data between April 1995 to March 2000, again the statistics are obtained from recursive estimates and one-step ahead forecasts. These results suggest that for S&P 500 returns there is evidence of the nonlinear models outperforming the linear model on all of the forecast evaluation statistics, with the LSTARX model being preferred to the ESTARX model on two of the three statistics (this contrasts with the in-sample tests where the ESTARX model was preferred on all measures).[14]

A final exercise is to compare the forecasting accuracy of the two nonlinear models with the linear model, and to consider whether the associated forecast errors are significantly

[14] The in-sample results for the DJIA are similar to those reported for the S&P 500 returns except on the mean absolute percentage error where the linear is preferred. However, for DJIA returns the linear model is preferred on all three out-of-sample forecasts performance statistics.

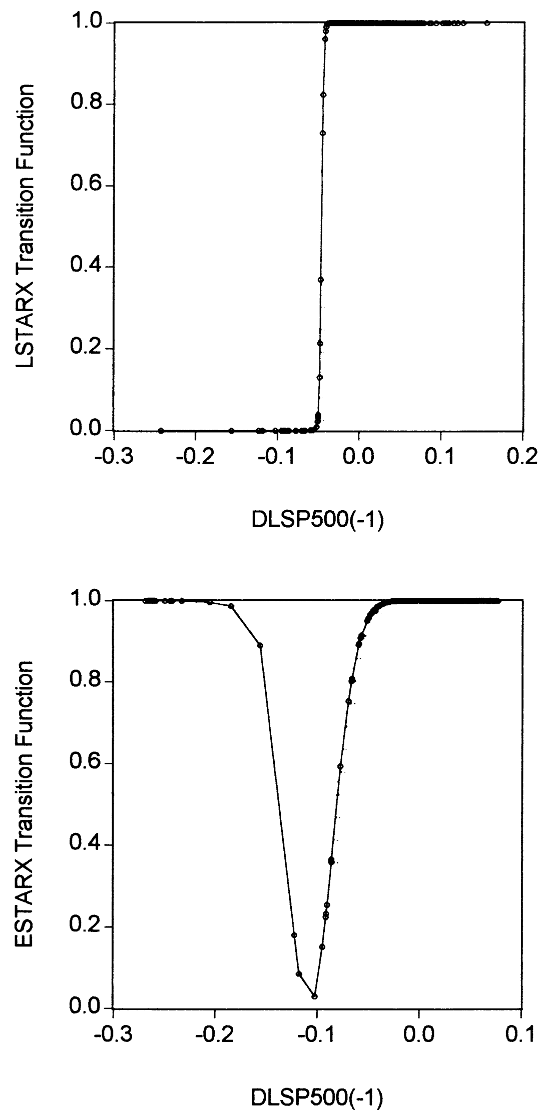

Fig. 2. STARX transition functions.

different, this is particularly pertinent given the small differences between the forecast statistics in Table 3. In order to conduct this analysis we perform the equality of forecast accuracy test of Diebold and Mariano (1995).[15] Given two forecast errors $\{e_{it}\}\{e_{jt}\}$, Diebold and Mariano define the forecast loss differential as $d_t = e_{it} - e_{jt}$ where a test of

[15] We are grateful to an anonymous referee for drawing our attention to this test.

Table 3
In-sample goodness of fit and out-of-sample forecast performance

Measures of in-sample fit				Measures of out-of-sample predictive power		
Model	RMSE	MAE	MA%E	RMSE	MAE	MA%E
Linear	.0423	.0316	218.7990	.0434	.0347	107.2348
LSTARX	.0412	.0309	191.2423	.0431*	.0344*	106.0918
ESTARX	.0408*	.0307*	177.8957*	.0433	.0345	104.6786*

* Lowest statistic.

equal forecast accuracy, i.e., $E[e_{it}]=E[e_{jt}]$ is equivalent to the mean of the loss differential being zero. Following Diebold and Mariano, the large sample mean of the loss differential, $\bar{d}$ is approximately normally distributed with mean μ and variance $2\pi f_d(0)/T$, where the large sample $N(0,1)$ statistic for the null hypothesis of equal forecasting accuracy is given by Eq. (12):

$$S_1 = \frac{d}{\sqrt{\frac{2\pi\hat{f}_d(0)}{T}}} \tag{12}$$

where $\hat{f}_d(0)$ is consistent estimator of $f_d(0)$. Given the potential for serial correlation in the loss differential, d, a consistent estimator of $2\pi f_d(0)$ is obtained by taking a weighted sum of the sample autocovariances (Eq. (13)):

$$2\pi\hat{f}_d(0) = \sum_{\tau=-(T-1)}^{(T-1)} 1\left(\frac{\tau}{S(T)}\right)\hat{\gamma}_d(\tau) \tag{13}$$

where $1(\tau/S(T))$ is the lag window, $S(T)$ the truncation lag and $\hat{\gamma}_d(\tau)$ the sample autocovariances at displacement τ. A variety of lag window and truncation lag choices are available, following Diebold and Mariano the uniform or rectangular lag window is selected while several truncation lags are considered, these being 5, 10 and 20.[16]

Fig. 3 presents the loss function for d_{LSTARX} and d_{ESTARX} where both are defined as the linear forecast error minus the nonlinear forecast error. Thus, positive loss differentials indicate a smaller forecast error for the nonlinear model, which becomes more apparent in the latter half of the forecast sample. The forecast accuracy equality test statistics for null hypothesis of d_{LSTARX} being equal to zero are 4.33, 4.66 and 19.35 for truncation lags of 5, 10 and 20, respectively, while the tests statistics for d_{LSTARX} are 3.13, 2.77 and 8.65, respectively. Thus, all statistics are significant at the 1% significance level. Therefore,

[16] Diebold and Mariano consider a $(k-1)$ truncation lag for the k step-ahead forecast errors; however, this is not feasible with the one step-ahead forecasts considered here.

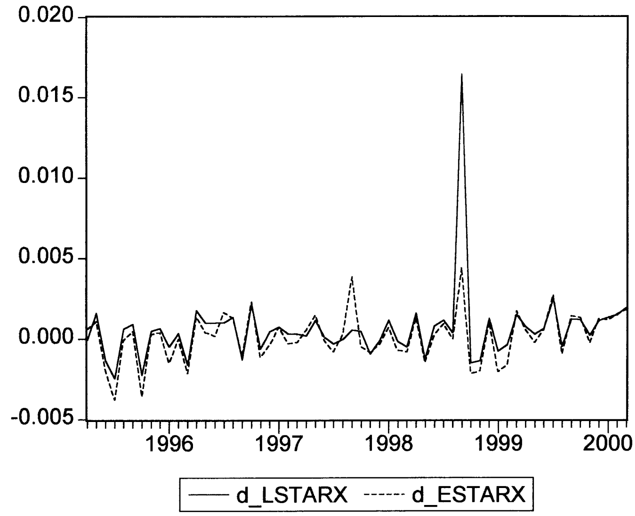

Fig. 3. Forecast loss differential.

despite the relatively small difference in standard forecast performance statistics, the nonlinear models do provide additional forecasting accuracy over the linear model.

4. Summary and conclusion

This paper has tested for evidence of a nonlinear relationship between stock market returns and macroeconomic and financial variables, and whether this nonlinearity can be exploited to improve forecasts of returns. Recent research has reported that stock market returns can be predicted using various lagged exogenous variables, such as interest rates and output measures. However, the majority of work in this area has presupposed a linear functional form, this paper has sought to reevaluate that hypothesis, and examine whether a nonlinear form can provide better forecasts. A linear model is initially estimated, with the result that various interest rates and macroeconomic variables provide some predictive power for S&P 500 returns. These variables are then used to investigate whether a nonlinear relationship exists using model-free nonparametric methods. The results of this exercise suggest that a nonlinear relationship does indeed exist between returns and interest rates, but not between returns and the macroeconomic series. The nonparametric plots suggest some possible threshold effect between returns and interest rates and thus we proceed to estimate a STAR-type model, termed STARX. This model supports the nonparametric results in that a nonlinear form is successfully estimated, with the interest rate series appearing in both regimes, while the macroeconomic series only appears in one regime. Tests measuring the in-sample goodness of fit support the nonlinear model over the linear model, similarly the results for the out-of-sample forecasting performance select the nonlinear model over the linear alternative, although the forecasting gain is marginal. Finally, tests of predictive accuracy suggest that although this gain may be marginal it is statistically significant.

In sum, this paper has reported some evidence of nonlinear predictability of stock market returns using financial variables, more specifically interest rates. Both the nonparametric plots and the estimated STARX models suggest the presence of nonlinear behaviour that may be exploited in forecasting exercises. Results suggest that the nonlinear models outperform the linear model both in-sample and out-of-sample, although the forecast gain is marginal, it thus remains an avenue for further research to see if alternative nonlinear forms can provide a better forecasting performance.

Acknowledgments

We acknowledge the helpful suggestions of the editor and three anonymous referees on an earlier version of this paper. All remaining errors are our own.

References

Abhyankar, A., Copeland, L. S., & Wong, W. (1997). Uncovering nonlinear structure in real-time stock market indices. *Journal of Business and Economics Statistics*, *15*, 1–14.

Balvers, R. J., Cosimano, T. F., & McDonald, B. (1990). Predicting stock returns in an efficient market. *Journal of Finance*, *45*, 1109–1128.

Breen, W., Glosten, L. R., & Jagannathan, R. (1990). Predictable variations in stock index returns. *Journal of Finance*, *44*, 1177–1189.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics*, *18*, 3–37.

Campbell, J. Y., & Hamao, Y. (1992). Predictable stock returns in the United States and Japan: a study of long-term capital integration. *Journal of Finance*, *47*, 43–67.

Chan, K., & Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, *7*, 179–194.

Cochrane, J. H. (1991). Production-bases asset pricing and the link between stock return and economic fluctuations. *Journal of Finance*, *46*, 209–238.

Delgado, M. A., & Robinson, P. M. (1992). Nonparametric and semiparametric methods for economic research. *Journal of Economic Surveys*, *6*, 201–249.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–265.

Epanechnikov, V. A. (1969). Nonparametric estimation of a multivariate probability density. *Theory of Probability and its Applications*, *14*, 153–158.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, *25*, 23–49.

Ferson, W. E., & Harvey, C. R. (1993). The risk and predictability of international equity returns. *Review of Financial Studies*, *6*, 527–566.

French, K. R., Schwert, G. S., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, *19*, 3–30.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, *48*, 1779–1801.

Granger, C. W. J., & Teräsvirta, T. (1993). Modelling nonlinear economic relationships. Oxford: Oxford University Press.

Haggan, V., & Ozaki, T. (1981). Modelling non-linear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, *68*, 189–196.

Härdle, W. (1990). *Applied nonparametric regression.* Cambridge: Cambridge University Press.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, *17*, 357–390.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, *9*, 141–142.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock market volatility. *Journal of Econometrics*, *45*, 267–290.

Perez-Quiros, G., & Timmermann, A. (2000). Firm size and cyclical variations in stock returns. *Journal of Finance*, *55*, 1229–1262.

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: robustness and economic significance. *Journal of Finance*, *50*, 1201–1228.

Pesaran, M. H., & Timmermann, A. (2000). A recursive modelling approach to predicting UK stock returns. *Economic Journal*, *110*, 159–191.

Poterba, J. M., & Summers, L. H. (1988). Mean reversion in stock prices. *Journal of Financial Economics*, *22*, 27–59.

Prakasa Rao, B. L. S. (1983). Nonparametric functional estimation. Orlando: Academic Press.

Qi, M. (1999). Nonlinear predictability of stock returns using financial and economic variables. *Journal of Business and Economic Statistics*, *17*, 419–429.

Robinson, P. M. (1983). Nonparametric estimators for time series. *Journal of Time Series Analysis*, *4*, 185–207.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of density function. *Annals of Mathematical Statistics*, *27*, 832–837.

Silverman, B. W. (1986). Density estimation of statistics and data analysis. London: Chapman & Hall.

Stone, C. J. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of Royal Statistical Society*, *36*, 111–147.

Stone, C. J. (1977). Consistent nonparametric regression (with discussion). *Annals of Statistics*, *5*, 595–645.

Stone, C. J. (1984). An asymptotically optimal window selection rule for kernel density estimates. *Annals of Statistics*, *12*, 1285–1297.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, *89*, 208–218.

Teräsvirta, T., & Anderson, H. M. (1992). Characterising nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics*, *7*, S119–S136.

Ullah, A. (1998). Non-parametric estimation of econometric functionals. *Canadian Journal of Economics*, *21*, 625–656.

Watson, G. S. (1964). Smooth regression analysis. *Sanikhya*, *26*, 175–184.