A discusión

ASYMMETRY IN THE LINK BETWEEN THE YIELD SPREAD AND INDUSTRIAL PRODUCTION. THRESHOLD EFFECTS AND FORECASTING*

Ivan Payá, Ioannis A. Venetis and David A. Peel**

WP-AD 2004-41

Corresponding author: Ivan Paya, Departamento Fundamentos Análisis Económico, Alicante University, 03080 Alicante, Spain. E-mail: ivanpaya@merlin.fae.ua.es Tel: +34 965903400 (ext 3221). Fax: +34 965903898

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

** Ivan Paya, Departamento Fundamentos Análisis Económico, Universidad de Alicante
I.A. Venetis: Centre of Planning and Economic Research (KEPE), Athens, Greece
D. A. Peel: Lancaster University Management School, United Kingdom

ASYMMETRY IN THE LINK BETWEEN THE YIELD SPREAD AND INDUSTRIAL PRODUCTION. THRESHOLD EFFECTS AND FORECASTING

Ivan Payá, I. A. Venetis and David A. Peel

ABSTRACT

We analyze the nonlinear behavior of the information content in the spread for future real economic activity. The spread linearly predicts one year ahead real growth in nine industrial production sectors of the US and four of the UK over the last forty years. However, recent investigations on the spread-real activity relation have questioned both its linear nature and its time-invariant framework. Our, in-sample, empirical evidence suggests that the spread real activity relationship exhibits asymmetries that allow for different predictive power of the spread when past spread values were above or below some threshold value. We then measure the out-of-sample forecast performance of the nonlinear model using predictive accuracy tests. The results show that significant improvement in forecasting accuracy, at least for one-step ahead forecasts, can be obtained over the linear model.

Key words: industrial production; yield spread; threshold model; forecasting; predictive accuracy

1 INTRODUCTION

Over the past few years researchers have discovered that differentials between long-term and short-term nominal interest rates (term spread) have exhibited a systematic correlation with movements in future economic activity in a number of countries and for a number of different time periods in the post-war period (see, e.g., Stock and Watson, 1989; Estrella and Hardouvelis, 1991; Plosser and Rouwenhorst, 1994; Estrella and Mishkin, 1997; Dotsey, 1998 and Galbraith and Tkacz, 2000). The empirical evidence reported to date suggests that the term spread is one of the most consistent predictors of future changes in real output and for horizons of a year or so ahead, has outperformed other potential leading indicators of output change, such as the money stock, stock prices or interest rates.

Recent theoretical work, by for instance Estrella (1998) and Estrella, Rodrigues and Schich (2000), Hamilton and Kim (2001), suggest a number of different ways by which a positive link between the term spread and future output change can emerge. Hamilton and Kim (2001) show that the contribution of the spread can be decomposed into the effect of expected future changes in short rates and the effect of the term premium where the respective contributions differ. Estrella (1998) derives the reduced form relationship between changes in real output and the spread from a simple linear structural model of the economy of the type specified by Fuhrer and Moore (1995).

Estrella (1998) demonstrates how the relationship between spread and real economic activity depends, in part, on the preferences of the policymaker between inflation and output deviations from target. In particular, when the authorities give higher weight to inflation the relationship between the spread and future real activity is weakened. This is because the impact of any future changes in expected inflation on future real activity is smaller. This is most easily seen by supposing that the authorities give zero weight to output stabilization. In these circumstances, expected inflation would be constant and the link between expected changes in future inflation and real output eliminated. In the latter context it is interesting that Ben Salem and Collard (2000) found that the empirical description of monetary policy by linear Taylor rules could be improved by employing a smooth transition autoregressive (STAR) model. A nonlinear Taylor rule accords with the hypothesis advanced by Orphanides and Wilcox (1996) who propose the so-called opportunistic approach to disinflation. In their model, the weights given to inflation and output stabilization by the monetary authorities are variable and dependent on the state of the economy, as measured by whether output and inflation are above or below target. In particular, policymakers will react more aggressively to bring inflation towards target whenever the inflation rate lies above or below certain threshold level. This particular threshold level would depend on various factors such as

the policymakers' preferences about output deviations from potential or the slope of the Phillips curve. As long as the inflation rate falls within the "acceptable" boundaries, the policymaker will concentrate on output stabilization. However, if inflation is either considered too high or too low, policymakers would attempt to bring it back to a certain level. This behaviour, due to the variable policy parameters, implies a nonlinear relationship between the spread and its forecasting ability over future output. Following this theoretical argument, we expect to find in empirical work that the spread has a stronger relationship with future output when the monetary authorities are relatively more concerned with output than with inflation. This would occur for relatively small previous values of the spread since they would be indicative of weaker current economic activity.[3]

Until recently one common feature of previous empirical work is that the relationship between aggregate output change and the spread has typically been modelled in a linear framework without investigating the possibility of asymmetric effects.

Galbraith and Tkacz (2000) report empirical evidence that is suggestive that the empirical relationship between the spread and aggregate output change may be nonlinear as the theoretical work suggests. They test for asymmetry in the predictive content of the spread in the form of a threshold effect. Their study finds evidence for the United States and Canada of the presence of an asymmetric impact of the conditional expectation of output growth depending on the position (above or below the threshold) of past spread values, though the reduced form they estimate includes other variables such as changes in government expenditure. Tkacz (2001) employs neural network models and documents the improved forecast accuracy (in terms of lower forecast errors) that can be achieved using nonlinear models to link the yield spread to aggregate output change. Venetis et al., (2001) examine the strength of the link between the yield spread and aggregate output change over the last forty years. This study confirms that threshold effects exist for a number of forecasting horizons affecting the power of the spread as a leading indicator in the case of the US, Canada and UK. It further shows a varying strength in the link with the spread-real activity relationship becoming weaker (or non-existent) after the adoption of inflation orientated policies by the respective Central Banks, something which is in line with the theoretical work of Estrella (1998).

Recently, Hansen (1997) has developed new statistical methods to aid in the estimation and specification of nonlinear threshold autoregressive (TAR) models. In this paper, we follow the methodology proposed by Hansen (1997) and we estimate and analyze the forecasting performance of threshold models of the yield spread-real activity relationship, relative to linear alternatives, employing data for several different industrial sectors in the US and UK. The out-of-sample predictive

---

[3]We also experimented in our empirical work with inflation or output change in the transition function but the results were superior employing the spread.

accuracy of different models is statistically tested using the Diebold and Marianno (1995) test. A number of other authors have examined the predictability of disaggregated output (Osborne et al., 2002 and references cited therein). However as yet non has considered a threshold model employing the theoretically appealing spread variable as the sole independent variable. We recognize from the theoretical work that the relationship linking the spread to aggregate real activity can differ between different economies dependent upon the policy regimes in operation and the different structural parameters between countries. Similarly any relationship between the spread and industrial production in different sectors of the economy may be expected to differ given varying response rates of output change to shocks. For instance, we may have different degrees of wage or price stickiness.

The rest of the paper is organized as follows. Section 2 presents the econometric methodology employed in the estimation of the nonlinear TAR models. Section 3 describes the data and provides empirical results for the linear and nonlinear estimations of the relationship between the term spread and future real economic activity for different production sectors in the US and UK in the post-war. Section 4 evaluates the out-of-sample forecast accuracy of both linear and nonlinear models. Finally, section 5 concludes.

2 ECONOMETRIC METHODS

We start by specifying linear leading indicator models for the annualized real industrial production growth over the next k months, $\Delta y_t^k = 100 \times \frac{12}{k} \times (y_{t+k} - y_t)$ where y_t is the logarithm of the underlying production series and we set the horizon $k = 12$. The particular dating of the spread variable, s_t, reflects its use as a leading indicator. As in many previous studies the basic model has the general form:

$$\Delta y_t^k = a_0 + a_1 s_{t-1} + u_t \tag{1}$$

where u_t represents a random disturbance term with mean zero and variance σ^2. The spread slope parameter represents an $a_1\%$ increase on the average $k-$ months forward growth to an 1% (100 basis points) increase in the spread.

Next, we consider the threshold autoregression[4] (TAR) model which will be compared with the linear model (1). Specifically

---

[4]The idea of approximating a general nonlinear autoregressive structure by a threshold autoregression with a small number of regimes is due to Tong (1983, 1990).

$$\Delta y_t^k = a_0 + a_1 s_{t-1} + u_t \quad \text{if } z_{t-d} \leq c \qquad (2)$$

$$\Delta y_t^k = b_0 + b_1 s_{t-1} + u_t \quad \text{if } z_{t-d} > c$$

The integer d is called the delay lag and typically it is unknown so it must be estimated. As we will shortly explain, the least-squares principle allows d to be estimated along with the other parameters. Parameter c is the "threshold" that distinguishes two regimes i) transition variable z_{t-d} is below c (lower regime) ii) transition variable z_{t-d} is above c (upper regime). Then, parameter vectors $\boldsymbol{\alpha} = (a_0, a_1)'$ and $\mathbf{b} = (b_0, b_1)'$ determine the total production growth response to changes in last period's spread.

If the threshold value, c, were known, then to test for threshold behavior all one needs is to test the hypothesis $H_0 : \boldsymbol{\alpha} = \mathbf{b}$. Unfortunately, the threshold value is typically unknown and, under the null hypothesis, parameter c is not identified[5]. The second difficult statistical issue associated with TAR models is the sampling distribution of the threshold estimate. Our model specification and inference will closely follow Hansen (1997) who a) provides a bootstrap procedure to test H_0, b) develops an approximation to the sampling distribution of the threshold estimator free of nuisance parameters and c) develops a statistical technique that allows confidence interval construction for c. In particular, we write TAR model (2) compactly as,

$$\Delta y_t^k = x_t(c)' \theta + u_t \qquad (3)$$

where $x_t(c) = (x_t' \mathbf{1}\{z_{t-d} \leq c\} \; x_t' \mathbf{1}\{z_{t-d} > c\})'$ with $x_t = (1, s_{t-1})'$, $\mathbf{1}\{.\}$ the indicator function and $\theta = (\boldsymbol{\alpha}', \mathbf{b}')'$. For a given value of c the least squares (LS) estimate of θ is

$$\hat{\theta}(c) = \left(\sum x_t(c) x_t(c)'\right)^{-1} \left(\sum x_t(c) \Delta y_t^k\right) \qquad (4)$$

with LS residuals $\hat{\xi}(c)_t$ and LS residual variance $\sigma_T^2(c) = (1/T) \sum_{t=1}^{T} \hat{u}^2(c)_t$. Then the LS estimate of c is the value,

$$\hat{c} = \arg\min_{c \in C} \sigma_T^2(c) \qquad (5)$$

where C is an interval (usually trimmed) that covers the sample range of the transition variable. Problem (5) can be solved by a direct search over C. The LS estimate of θ is then, $\hat{\theta} = \hat{\theta}(\hat{c})$.

---

[5] Nothing can be learned about c from the data when the null hypothesis is true.

Furthermore, the LS principle allows us to estimate the, typically, unknown value of d by extending problem (5) to a search across the discrete space $[1, \bar{d}]$.

The hypothesis $H_0 : \boldsymbol{\alpha} = \mathbf{b}$ is tested as follows: Let $\{e_t\}_{t=1}^T$ be an $i.i.d$ sequence of $N(0,1)$ draws. Regress e_t on x_t to obtain the residual variance $\hat{\sigma}_T^2$ and on $x_t(c)$ to obtain $\hat{\sigma}_T^2(c)$ and compute $F(c) = T\left(\frac{\hat{\sigma}_T^2 - \hat{\sigma}_T^2(c)}{\hat{\sigma}_T^2(c)}\right)$. Then compute $F = \sup_{c \in C} F(c)$. Repeat the procedure n times (we set $n = 1000$) and the asymptotic p-value of the test is given by the percentage of samples for which F exceeds the observed F_T (is based on calculations using the observed data sample).

Finally, Hansen (1997) provides critical values and a method to construct asymptotically valid confidence intervals. Estimate the model using the actual data for a set of values of c in the range C and in each case calculate the likelihood ratio statistic $LR(c)$ for that value of c against the value of the likelihood obtained by unrestricted LS, that is, $LR(c) = T\left(\frac{\hat{\sigma}_T^2(c) - \hat{\sigma}_T^2(\hat{c})}{\hat{\sigma}_T^2(\hat{c})}\right)$. Notice that for $c = \hat{c}$ we get $LR(c) = 0$. Then plot $LR(c)$ against c and draw a flat line that corresponds to the β-level critical value $c^*(\beta)$ given in Hansen (1997, table 1, p.5). For $\beta = 5\%$ we have $c^*(\beta) = 7.35$. The confidence interval LR^c is given by $LR^c = \{c : LR(c) \leq c^*(\beta)\}$.

3 ECONOMETRIC RESULTS

We apply model (2) to nine U.S monthly industrial production series for the period 1960.01-1999.03 and to four U.K. monthly real industrial production series. The yield spread is simply the difference between the long- and short-term nominal interest rates at any point in time (10 year government bond yield and 3-month treasury bill rate respectively). All data were obtained from the DATAS-TREAM database. Industrial production series for U.S begin with an "s" and for U.K. with a "k". In particular, sip: industrial production, scg: consumer goods, sdg: durable goods, sig: investment goods, sim: intermediate goods, smn: manufacturing goods, snd: non-durable goods, srm: raw materials, ssc: supply goods. In the U.K. case we have, kip:1960.01-1999.02, kcg:1968.01-1995.04, kig:1968.01-1999.02, and kim:1968.01-1999.02.

Given that the monthly spread is too volatile and given our focus on the asymmetry of the yield-output link, we follow Galbraith and Tkacz (2000) and we set z_{t-d} to be a function (moving average) of past spread values. Four possible specifications of the threshold variable were employed, $z_{q,t-i} = \frac{1}{(q-i+1)} \sum_{j=i}^{q} s_{t-j}$ for $(i,q) = ((1,6),(1,12),(6,12),(6,18))$. In all cases the transition between regimes depends on a moving average of past spread values. The search interval was set to be $C = [\underline{c}, \bar{c}] = [\min\{z_{q,t-i}\} + 0.5, \max\{z_{q,t-i}\} - 0.5]$. Notice that in order to "gain" observations, we arbitrary add or subtract 0.5 to construct the boundaries $\underline{c}, \bar{c}$ and then we divided C into 200 discrete

points. We experienced no problems when solving problem (5) although an alternative approach would be to discard 15% of the observations in both ends of $z_{q,t-i}$. Our final choice of $z_{q,t-i}$ was based on the minimum residual variance and on the maximum F statistic.

Our results are reported in tables 1 (linear model) and 2 (nonlinear model). We will not comment extensively on the linear model as it has been widely applied and it was adopted in our analysis as a comparing and contrasting tool. Nevertheless, our findings are in accordance with the existing literature with a_1 being statistically significant and $\hat{a}_1$ being positive and economically significant. For example in the U.S case an 1% increase in the spread last month over the sample period implies on average 1.377% increase in real industrial production over the following 12 months. Our results for the U.K. series are again not surprising. The slope parameter is less economically significant at around 0.65 for most of the series but still statistically significant. Plosser and Rouwenhorst (1994) suggested that for U.K. "*the information in the nominal term structure about real activity is obscured by variations in inflationary expectations*" since the U.K. experienced relatively high and volatile inflation.

The nonlinear estimations are reported in Table 2. The most relevant finding is the evidence of a threshold effect in the leading indicator equation that implies different predictive ability depending on whether the spread is above or below such a threshold. The F test uniformly and strongly rejects the null of equal parameters above and below the threshold for all series. This result confirms the asymmetry of the predicting power of the spread pointed out earlier in the literature regarding the real GDP-yield spread link (see Galbraith and Tkacz, 2000; and Venetis et al., 2001). In the case of the US, most of the estimated threshold values are around 1.4 percent, except for consumption and non-durable goods where the threshold is almost half. The asymptotic 95 percent confidence intervals are not particularly wide and the estimated thresholds are situated usually in the middle of the interval reinforcing the robustness of the point estimates. Only for the durable goods case we obtain a wide confidence interval pointing to some uncertainty regarding the point estimate.

A particular case is the investment goods sector where the estimated threshold is negative. This finding could be due to the fact that the investment goods series have the highest volatility and large negative values for rates of change of investment goods are found. We interpret such volatility in the sense that investment goods proportionally over-react to economic expansions and slowdowns. Nevertheless, the estimated slope parameters are significant either above or below the threshold meaning that the spread correctly signals both expansions and recessions in investment goods.

For the rest of the U.S series, the estimated slope parameter when the spread is below $\hat{c}$ is always statistically and economically significant but when past spread values move above $\hat{c}$ the predictive ability of the spread ceases to exist. However, for most of the U.S series around 60% of

the observations lie below, (or equal) the threshold implying a consistent predictive ability over production along time.

The ratio $\hat{\sigma}_{NL}/\hat{\sigma}_L$ where $\hat{\sigma}_{NL}^2$, $\hat{\sigma}_L^2$ are the residual variances of the nonlinear and linear models respectively, provides a measure of in-sample fit comparison for the alternative models. Given that it is less than one it implies better fit and as a rule of thumb if it is less than 0.90 suggests that nonlinearity explains more than just exceptional observations in the sample. The non-linear models appear to outperform the linear ones in all cases. For the US the ratio of standard errors is less than 0.90 except for the consumption goods series. For the durable and supply goods models it is as low as 68%.

A different story emerges when we examine the U.K. results in table 2. The U.K. results differ not only quantitatively but also qualitatively. The ratio $\hat{\sigma}_{NL}/\hat{\sigma}_L$ is in all cases above or equal to 0.90 suggesting the nonlinear model is called to "explain" only exceptional circumstances. Notice that most of the observations lie under the estimate threshold value which is 1.29 for kcg and 1.89 for kip and kim with the slope parameter being statistically insignificant. In the upper regime, the spread slope parameter is statistically and economically significant. For the kim series $\hat{b}_1$ is as high as 3.46. These estimates are in contrast with the results obtained for the U.S series. This difference suggests that the asymmetry in the link between yield spread and growth is less intense in the case of the UK over the sample period. This is supported by the fact that the spread is significant in both the upper and lower regime (except in kim) and that the $\hat{\sigma}_{NL}/\hat{\sigma}_L$ ratio does not go below 0.90. The spread in the UK is then capable of predicting large economic expansions as well as slowdowns. Only for the investment goods, kig, we obtained results comparable with the U.S ones. For the kig model, the threshold is low at 0.36. In the lower regime spread has predictive power with a coefficient of 2.13. In the upper regime, the predictive power disappears and we cannot reject the hypothesis $b_1 = 0$. Again, the low threshold value found in this series and its volatility plays a role in the ambiguity of the results found in the investment goods sector.

4 COMPARING FORECASTING ACCURACY

In this section we provide a forecasting experiment. Our purpose is to compare the out-of-sample forecasts of the two alternative models (1) and (2). We have already discussed in-sample fit issues with the nonlinear model clearly dominating the linear one at least in terms of residual variance minimization in the U.S case although the U.K. results are difficult to interpret.

For the U.S samples, we retain $m = 183$ observations for out-of-sample forecasting. In particular

9

we will calculate one step ahead forecasts for the period 1983.01 onwards[6]. In order to reduce parameter uncertainty, both linear and nonlinear models were re-estimated for each observation added to the sample after the date 1983.01 and the forecasts are based on the "updated" parameter estimates. For the U.K. samples, we also retain all observations after 1983.01 for re-estimation of the leading indicator models and forecasting.

The forecast evaluation will be based on two different criteria. Let $\hat{y}_{T_0+j|T_0+j-1}$ denote the forecast of Δy_t^k, at time $T_0 + j$ given information up until time $T_0 + j - 1$ where, for example, T_0 corresponds to 1983.01 in the U.S forecasting experiment. The mean squared prediction error is given by

$$MSPE = \frac{1}{m} \sum_{j=1}^{m} \left(\hat{y}_{T_0+j|T_0+j-1} - y_{T_0+j} \right)^2 \tag{6}$$

and the mean absolute deviation error is given by,

$$MAPE = \frac{1}{m} \sum_{j=1}^{m} \left| \hat{y}_{T_0+j|T_0+j-1} - y_{T_0+j} \right| \tag{7}$$

Although these measures can provide a first indication on whether a certain model performs better on average (in terms of minimizing the respective loss function), they cannot provide statistical significance results on the difference of forecasts. In order to test the null hypothesis of equal forecasting accuracy we will employ the following two tests. Consider the sample path $\{d_j\}_1^m$ of a loss differential series, that is, d_j is a function of the difference of forecast errors produced by two different models. For our application we chose the quadratic function,

$$d_j = \left((\hat{y}_{T_0+j|T_0+j-1,NL} - y_{T_0+j}) - (\hat{y}_{T_0+j|T_0+j-1,L} - y_{T_0+j}) \right)^2 \tag{8}$$

although the absolute difference function did not qualitatively alter our results.

Following Diebold and Mariano (1995), the large sample studentized version of an exact finite sample test, "the sign test" is given by,

$$S = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \left(\mathbf{1}\{d_j > 0\} - 0.5 \right) \sim N(0,1) \tag{9}$$

The test compares the relative magnitudes of the prediction errors of the two models but a serially correlated loss differential is not allowed under the test's assumptions. Diebold and Mariano

---

[6] This date is close to the estimated sample break in the predicting power of the US spread over GDP found in Venetis et al. (2001).

(1995) propose the use of an asymptotic statistic which tests whether the average loss differential is significantly different from zero. Specifically they show that,

$$DM = \frac{\bar{d}}{\sqrt{\hat{\omega}/m}} \rightarrow_d N(0,1) \tag{10}$$

where $\hat{\omega}$ is an estimate of the long run covariance matrix of d_j.

When forecasts are constructed from ARMA(p,q) models then the infinite moving average representation produces h-step ahead prediction errors that are serially correlated up to order h-1. As a result $\hat{\omega} = \hat{\gamma}_0(d)$ when the DM test is applied to 1-step ahead forecasts. Nevertheless, our model is nonlinear and although $h = 1$ we found that in most cases d_j was strongly correlated. For example in the case of sip, $\hat{\gamma}_1(d) = 0.928, \hat{\gamma}_2(d) = 0.840, \hat{\gamma}_3(d) = 0.756, \ldots$. For that reason we set $\hat{\omega}$ equal to the long run covariance matrix of d_j and the Newey-West (1987) estimate is adopted with $\hat{\omega}_{NW} = \hat{\gamma}_0(d) + 2 \sum_{v=1}^{q} (1 - \frac{q}{v+1}) \hat{\gamma}_v(d)$. We chose the truncation lag q according to the correlation structure of d_j. As a rule of thumb we calculated the autocorrelation function of d_j and then we selected the value of q that reflected the number of lags with significant correlations. For example in the sip model we set $q = 7$.

Our results are summarized in table 3. With the exception of the U.K. series of intermediate goods, kim, the MSPE's and MAPE's produced by the nonlinear model are uniformly lower than their linear model counterparts. In the case of U.S, the DM statistic, at least for the 10% significance level, implies that for 7 out of 9 series the null hypothesis of equal predictive accuracy is rejected. On the contrary, but in accordance with their poor relative in-sample performance, the DM statistic for the U.K. series did not produce statistically significant results.

Figure 1 plots the actual (indicated by a plain line) and one step ahead nonlinear (line with pluses) and linear forecasts (line with boxes) for the three of the U.S production series having the largest, in absolute terms, DM statistics namely, scg, sim and snd. These plots illustrate the improvement in fit that occurs when the nonlinear threshold model is considered. With the exception of a (roughly) two year period that coincides with the latest U.S recession, the linear model forecasts consistently "overshoot" predicting one year ahead growth that is constantly above the actual one. On the contrary, the nonlinear forecasts move closely to the actual series and for many periods seem to track the mean of the actual growth rates.

Finally, it is worth commenting on the difference between the sign, "S", and DM statistics, the latter being subjected to correlation correction. In a number of cases, the correction significantly affects the magnitude of the statistics changing our decision relative to the null hypothesis. For example, in the sig series model we get $S = -4.36$ but $DM = -0.84$.

11

5 CONCLUSIONS

This paper analyses the asymmetry in the relationship between the yield spread and economic activity in the US and UK. In particular, linear and nonlinear (TAR) models present significant predicting effects over nine and four different industrial production sectors in the US and UK respectively from 1960 to 1999. The nonlinear models appear to outperform the linear ones in the in-sample estimations as the ratio between regression standard errors evidences. Nevertheless, the significance difference between the two models appears to be much stronger in the case of the US industrial sectors consistent with previous studies on the link between spread and aggregate rates of change of real output. Recursive out-of-sample forecasts are also implemented for all different sectors from 1983 onwards comparing again the linear and the nonlinear models. The Diebold-Mariano forecast accuracy tests, namely, the "sign-test" (S) and the DM test, exhibits a significance difference between the two competing models. The forecasts using the TAR model produces more accurate values of future industrial activity in most of the US industrial sectors. In the case of the UK the evidence is weaker suggesting that there is not a significant superior forecast performance of the nonlinear model. It is of course difficult to be certain of the rationale for this outcome. Perhaps a prime candidate, suggested by the theoretical models, is that the monetary regimes operating in the US have been far more stable(or predictable) than in the UK. Whilst Clarida et al. (2001) find that Taylor monetary rules have been relatively stable in the US over the post war period,Nelson (2000) finds that Taylor rules have exhibited great instability in the UK over a similar period. As a consequence estimation of any nonlinear(linear) relationship between real activity and the spread in the UK may require much closer attention to the regimes in operation, e.g. Conservative or Labour Governments and Independent Central Bank as a minimum. For economies, such as the US, with more stable monetary regimes our results suggest that 'a more simple nonlinear threshold relationship' between real activity and the spread may become more discernible and may help produce superior forecasts to the linear alternative.

	$\hat{a}_0$	$s.e(\hat{a}_0)$	$\hat{a}_1$	$s.e(\hat{a}_1)$	$\hat{\sigma}_L$
sip	2.75	0.45	1.37	0.23	3.95
scg	2.44	0.37	0.77	0.16	3.29
sdg	3.12	0.59	2.40	0.31	5.35
sig	3.37	0.68	1.45	0.34	5.77
sim	2.46	0.49	1.22	0.25	4.20
smn	2.89	0.48	1.59	0.25	4.25
snd	2.62	0.40	0.75	0.21	3.45
srm	2.77	0.56	1.81	0.30	4.93
ssc	1.60	0.64	1.96	0.33	5.46
kip	1.17	0.40	0.66	0.23	3.90
kcg	0.92	0.53	0.63	0.27	4.25
kig	0.49	0.68	0.62	0.34	5.30
kim	-0.53	0.54	0.78	0.25	5.14

Table 1. Ordinary least squares results for the linear model (1). In parentheses we report the Newey-West (1987) standard error estimates of the parameters.

	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}$	F	LR^c	$\%obs \leq \hat{c}$	$\hat{\sigma}_{NL}/\hat{\sigma}_L$	$z_{q,t-i}$
sip	2.32	−0.23	1.42	145.23	[1.18	0.67	0.86	$z_{12,t-1}$
	0.12	0.31		[0.000]	1.53]			
scg	1.57	0.29	0.87	76.23	[0.63	0.51	0.93	$z_{12,t-1}$
	0.12	0.22		[0.000]	1.11]			
sdg	1.77	0.55	1.49	30.62	[1.25	0.69	0.68	$z_{12,t-1}$
	0.12	0.34		[0.000]	1.69]			
sig	1.11	−0.71	−0.30	145.90	[−0.44	0.29	0.87	$z_{12,t-1}$
	0.25	0.24		[0.000]	−0.20]			
sim	2.23	−0.50	1.18	117.61	[0.90	0.60	0.89	$z_{12,t-1}$
	0.14	0.30		[0.000]	1.49]			
smn	1.57	0.22	1.49	54.99	[1.28	0.69	0.74	$z_{12,t-1}$
	0.11	0.29		[0.000]	1.69]			
snd	1.73	−0.25	0.70	113.05	[0.63	0.47	0.90	$z_{12,t-6}$
	0.12	0.19		[0.000]	1.28]			
srm	3.01	0.18	1.42	154.18	[1.28	0.67	0.85	$z_{12,t-1}$
	0.14	0.38		[0.000]	1.49]			
ssc	1.46	0.76	1.60	45.73	[1.46	0.75	0.68	$z_{6,t-1}$
	0.12	0.41		[0.000]	1.71]			
kip	0.42	2.38	1.89	42.83	[1.74	0.79	0.93	$z_{18,t-6}$
	0.10	0.28		[0.000]	2.15]			
kcg	0.38	1.21	1.29	29.42	[1.26	0.65	0.94	$z_{18,t-6}$
	0.15	0.24		[0.000]	1.52]			
kig	2.13	−0.31	0.36	57.39	[0.03	0.33	0.91	$z_{12,t-6}$
	0.25	0.20		[0.000]	0.53]			
kim	0.19	3.46	1.89	80.32	[1.52	0.74	0.90	$z_{18,t-6}$
	0.14	0.36		[0.000]	1.93]			

Table 2. Least squares results for model (2). $\hat{c}$ is the estimated threshold parameter.In parentheses we report standard errors.F is the statistic to test the null $H_0 : \alpha = b$ using the bootstrap procedure of Hansen(1997) Squared brackets enclose p-values for the F statistic.LR^c denotes the 95% asymptotic confidence interval for c calculated as in Hansen (1997). $\%obs \leq \hat{c}$ denotes the percentage of observations below or equal to the threshold

	MSPE(L)	MSPE(NL)	MAPE(L)	MAPE(NL)	S	DM
sip	17.89	9.85	3.48	2.35	−2.44	−1.83
scg	9.90	6.70	2.74	1.98	−4.21	−2.20
sdg	17.96	15.97	3.33	3.07	−0.52	−1.19
sig	32.58	28.10	4.59	3.87	−4.36	−0.84
sim	22.02	14.25	4.02	2.82	−3.03	−2.08
smn	10.85	8.50	2.71	2.17	−2.43	−1.69
snd	13.50	8.19	3.22	2.34	−4.36	−2.14
srm	24.07	12.92	3.90	2.84	−1.40	−1.65
ssc	20.55	17.84	3.61	3.11	−1.70	−1.95
kip	7.67	7.11	2.25	2.14	−3.11	−0.98
kcg	10.65	10.21	2.68	2.62	−1.20	−0.27
kig	24.06	20.04	3.77	3.53	−3.40	−1.43
kim	16.06	16.39	3.04	3.09	0.88	0.16

Table 3. One step ahead forecast comparison. The S and DM tests are $N(0,1)$ distributed. The null hypothesis is equal forecasting accuracy for the linear and nonlinear model forecasts.

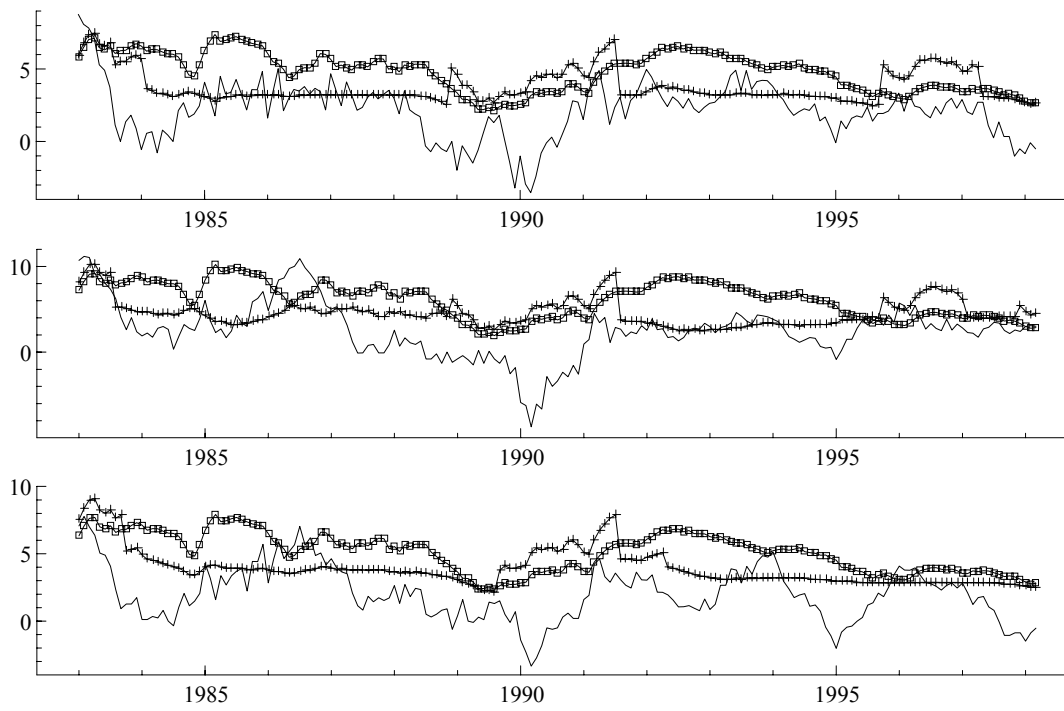In parentheses (L), (NL) denote linear and nonlinear models respectively.

Figure 1. U.S. Observed one year ahead growth series (line) and nonlinear line with "plus" symbols) and linear (line with "box" symbols) model one-step ahead forecasts. Top: scg, middle: sim, bottom: snd

REFERENCES

Bec F, Ben Salem M, Collard F. 2000. Nonlinear Economic Policies: Pitfalls in the Lucas Critique Empirical Counterpart. Paper No 1401 in *Econometric Society World Congress 2000* Contributed Papers

Diebold FX, Mariano S. 1995. Comparing predictive accuracy. *Journal of Business and Economics Statistics* **13**: 253-263.

Dotsey, M, 1998. "The predictive content of the interest rate term spread for future economic growth". *Federal Reserve Bank of Richmond Economic Quarterly* **84**: 31-51.

Estrella A, Hardouvelis GA. 1991. The term structure as a predictor of real economic activity. *The Journal of Finance* **46**: 555-576.

Estrella A, Mishkin FS. 1997. The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review* **41**: 1375-1401.

Estrella A. 1998**.** Monetary policy and the predictive power of the term structure of interest rates. Federal Reserve Bank of New York, November 1998.

Estrella A, Rodrigues, AP, Schich S. 2000. How stable is the predictive power of the yield curve? Evidence from Germany and the United States. *Federal Reserve Bank of New York*, September 2000.

Franses PH, van Dijk D. 2000. Non-linear time series models in empirical finance. Cambridge University Press.

Fuhrer JC, Moore GR. 1995. Monetary Policy Trade-offs and the Correlation between Nominal Interest Rates and Real Output. *American Economic Review*, **85**, 219-239.

Galbraith JW, Tkacz G. 2000. Testing for Asymmetry in the link Between the Yield Spread and Output in the G-7 Countries. *Journal of International Money and Finance* **19**: 657-672.

Hamilton JD, Kim, DH. 2001. A Re-Examination of the Predictability of Economic Activity using the Yield Spread. *Journal of Money, Credit and Banking* **34**: 491-520.

Hansen BE. 1997. Inference in TAR Models. *Studies in Nonlinear Dynamics and Econometrics* **2**: 1-14.

Nelson ,E 2000.UK monetary policy 1972-97: a guide using Taylor rules, Bank of England Working Papers

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

Orphanides A, Wilcox, DW. 1996. The Opportunistic Approach to Disinflation. No 96-24 working paper in Finance and Economics Discussion Series from *Board of Governors of the Federal Reserve System, U.S.*

Osborne D, Simpson P, Sensier M. 2002. Forecasting UK Industrial Production over the Business Cycle. *University of Manchester mimeo. Forthcoming in the Journal of Forecasting.*

Plosser CI, Rouwenhorst KG. 1994. International Term Structures and Real Economic Growth. *Journal of Monetary Economics* **33**: 133-155.

Stock JH, Watson, MW. 1989. New Indexes of Coincident and Leading Indicators, *NBER Macroeconomics Annual*, 351-393.

Tkacz, G. 2001. Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting*, **17**, 57-69

Tong H, 1983. Threshold Models in Non-linear Time Series Analysis: Lecture Notes in Statistics 21. Berlin: Springer-Verlag.

Tong H, 1990. Non-Linear Time Series: A Dynamical System Approach. New York: Oxford University Press.

Venetis, IA, Paya, I, Peel, DA. 2001. Re-examination of the predictability of economic activity using the yield spread: A nonlinear approach. I*nternational Review of Economics and Finance*, *forthcoming.*