A discusión

PREDICTING REAL GROWTH AND THE PROBABILITY OF RECESSION IN THE EURO AREA USING THE YIELD SPREAD*

Agustín Duarte, Ioannis A. Venetis and Ivan Payá**

WP-AD 2004-31

Corresponding autor: I. Payá, Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, Campus de San Vicente, 03080 Alicante, Spain.
Email: ivanpaya@merlin.fae.ua.es

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.

** A. Duarte: Department of Applied Economics and Economic Policy, University of Alicante, Spain
I.A. Venetis: Centre of Planning and Economic Research (KEPE), Hippokratous 22, Athens 106 80, Greece
I. Payá: Departamento Fundamentos Análisis Económico, University of Alicante, 03080 Alicante, Spain.

PREDICTING REAL GROWTH AND THE PROBABILITY OF RECESSION IN THE EURO AREA USING THE YIELD SPREAD

Agustín Duarte, Ioannis A. Venetis and Ivan Payá

ABSTRACT

Although the spread has been established as a leading indicator of economic activity, recent studies on US and EU countries have documented, theoretically and empirically, that the term spread-output growth relationship may not be stable over time and it may be subjected to nonlinearities. Using aggregate data for the Euro area over the period 1970:1 - 2000:4, we applied linear regression as well as nonlinear models to examine the predictive accuracy of the term spread-output growth relationship. Our results confirm the ability of the yield curve as a leading indicator. Moreover, significant nonlinearity with respect to time and past annual growth is detected outperforming the linear model in out-of-sample forecasts of one-year-ahead annual growth. Furthermore probit models that use the EMU and US yield spreads are successful in predicting EMU recessions.

1 Introduction

During the last two decades, the literature has accumulated an important volume of studies on the information content of the yield curve, the difference between a long term interest rate and a short-term interest rate, as an indicator to assess economic conditions and to predict real economic activity[2]. Intuitively, a positive sloped yield curve is associated with an increase in economic activity, in the rate of inflation and of short-term interest rates. On the contrary, a negative sloped yield curve is associated with a future decrease in the values of these variables[3]. The information content of the yield curve has been the object of countless theoretical and empirical work. The bulk of the studies have concentrated on the US economy (see e.g. Modigliani and Sutch, 1966; Turnovsky, 1989; Stock and Watson, 1989; Estrella and Hardouvelis, 1991) while others have also investigated some major European countries (see e.g. Plosser and Rouwenhurst, 1994; Bernard and Gerlach, 1996; Davis and Fagan, 1997, Estrella and Mishkin, 1997, and Berk and Van Bergeijk, 2000).

Although a number of empirical studies have documented the relevance of the yield curve to predict recessions and future output growth, more recently, the stability and predictive power of these relationships have been questioned both on theoretical and empirical grounds[4]. As some empirical studies have recently pointed out, the yield curve has been losing predicting power since the late 1980s. A classical example of this fact is its failure to predict the 1990-91 recession in the US[5]. Other empirical evidence, however, indicates that the prediction with models that use binary indicators of recession or expansions are more successful and stable than continuous ones (Estrella, Rodriguez and Schich, 2000).

From a theoretical point of view, Estrella (1998), and Hamilton and Kim (2002) have studied the above mentioned arguments. Following the model of Fuhrer and Moore (1995), Estrella develops a dynamic rational expectation model to explore the specific circumstances under which the yield curve has predictive power over some macroeconomic variables such as real output and inflation and 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 rela-

---

[2] The terms "yield curve", "term structure of interest rates", or "term spread" will be used interchangeably.

[3] Formally, these results follow from the combination of the Fischer equation on the one hand and the expectation theory of the yield curve on the other hand (Modigliani and Sutch, 1966), or alternatively from the standard IS-LM model for a small open economy (Blanchard and Fischer, 1989, p. 536).

[4] In a European context, Davis and Fagan (1997) study the usefulness of different financial spreads as indicators of future output growth. The yield curve provided the best within-sample results. However, only three countries (Belgium, Denmark and UK) satisfied the conditions of significance, stability and improved out-of-sample forecasts.

[5] See Haubrich and Dombrosky, 1996; Dotsey, 1998; Friedman and Kuttner, 1998; Stock and Watson, 2001.

tionship 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. 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.

A notable feature of most empirical works is that the term spread-output growth relation has been modelled using a linear framework and little attention has been given to the possibility of asymmetric effects and time varying parameters. Indeed, as some recent studies have shown, based on data for the US and Canada, the term spread-output relation might not be linear and its predictive content might also have asymmetric effects, as measured in terms of a threshold on the conditional expectation of output growth (Galbraith and Tkacz, 2000)[6]. In line with these arguments, Venetis et al. (2003) study the predictive power and stability of the spread-output relationship with data from the US, UK and Canada using non-linear autoregression models that can accommodate regime switching type non-linear behavior and situations of time varying parameters. A significant result of this work is that the spread-output growth relationship is stronger when past values of the term spread do not exceed a positive value of the threshold.

Cross country empirical evidence, including a wide sample of European countries, have also confirmed the usefulness of the yield curve to predict the probability of a recession and of future real economic activity (Bernard and Gerlach, 1996; Estrella and Mishkin, 1997; and Sensier et al., 2002). Those works have also demonstrated that the term spread-output growth relationship is stronger in countries where monetary policy is independent than in those countries that peg their currencies to a bigger commercial partner. This result is of particular relevance for the countries participating in the Euro area. As it is known, after the "Single European Act" agreement, signed in February 1986, most EU countries have achieved a high degree of nominal and real convergence under a quasi-fix exchange rate system. Under these circumstances, it should be admitted that the information content of the yield curve for individual countries could be very useful for both private investors and the European Central Bank[7]. However, it should also be emphasized that although this valuable information is necessary, it is by no means sufficient for the aggregate Euro area. To the extent that in some of the EU countries the yield curve might have little influence on output, it is important to know the information content of the yield curve using aggregate data.

The purpose of this paper is to provide new evidence on the information content of the yield curve in the Euro area. In particular, our aim is twofold. First to explore the stability as well as the predictive power of the term spread-output relationship. Contrary to previous studies, this paper builds on recent non-linear

---

[6]In the case of Canada, Tkacz (2001) uses neural network models and shows that non-linear models have smaller prediction errors than linear models.

[7]Recall that the yield curve has been included as part of the second "pillar" in the ECB Monetary Policy Strategy along with a number of alternative indicators (See: ECB, Monthly Report. January 1999, p.49).

econometric techniques and uses quarterly aggregate data over the period 1970:1 - 2000:4. Through the introduction of a threshold variable, we can capture potential non-linearities in the relationship within a tractable and intuitively appealing parametric model. Second, to examine the ability of the yield curve to predict the probability of a future recession. Recent analysis, in a multi-equation context questioning the issues of structural breaks, nonlinearity and the predictive ability of the spread has been undertaken by Galvao (2003) and Anderson and Vahid (2000). Multi-equation models provide a generalization for the single equation procedures we employ.

The rest of this paper is organized as follows. Section two focuses on the strength of the relationship between the yield curve and real economic activity. This section introduces the econometric methodology as regards stability tests and threshold models. The idea is to allow for nonlinearities in the conditional expectation function without over-parameterization. Then, empirical results and prediction from alternative specifications (linear and non-linear) are presented. In section three we apply probit models in order to predict the likelihood of a future recession and we discuss the major results. Section four concludes.

2   The yield curve and future real activity in the Euro area.

2.1   The linear model

We use aggregate quarterly data for the Euro area over the period 1970:1-2000:4[8]. In particular, real GDP, 10-year government bond yield and the 3-month interest rate on deposits. The annual real GDP growth and the spread series are plotted in Figures 1(a) and 1(b). We start by specifying linear leading indicator models for the annualized real GDP growth over the next k quarters, $\Delta y_t^k = 100 \times \frac{4}{k} \times (y_{t+k} - y_t)$ where y_t is the logarithm of the Euroland real GDP series and we set the horizon $k = 1, 2, 3, 4, 5, 6, 8, 12$. 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:

$$\Delta y_t^k = a_0 + a_1 s_{t-l} + u_t \qquad (1)$$

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 $k-$ quarters ahead growth to an 1% (100 basis points) increase in the spread. The specification of (1) was based on the Schwarz information criterion. We employed a number of lags $l = 1, ..., 8$ and combinations of lagged spread values as explanatory variables. Moreover, lagged real growth was also introduced into (1) but multivariate right hand side specifications did not improve upon (1). Table 1 reports the results of

[8]For a detailed description on data construction and data sources see Appendix 1.

fitting equation (1) for the Euro-land aggregated data set along with some structural break tests (last 4 rows) that will be explained in detail in the next subsection.

In parentheses, we report Newey and West (1987) corrected standard errors that take into account the moving average created by the overlapping of forecasting horizons as well as conditional heteroskedasticity. The second row reports the value of l that minimized the Schwarz criterion while the second row reports slope parameter estimates $\hat{a}_1$. At all horizons the estimates $\hat{a}_1$ are positively signed and highly significant. Our results can be easily compared with the ones found for individual European countries and the US in Estrella and Mishkin (1997) for similar time periods. The $\hat{a}_1$ coefficients in Table 1 are slightly higher than the ones reported in Estrella and Mishkin for France, Germany and Italy. However, compared with the slope parameter of the US, the Euro-land coefficient is slightly lower. In terms of the adjusted R^2 our results show much higher explanatory power for short horizons than the one found for individual countries. However, the highest R^2 values prevail at similar horizons, that is between three and six quarters ahead[9].

2.2 Nonlinear models

In this section we explore the stability of the linear model as well as two alternative nonlinear specifications with regard to the yield spread-output growth relationship in the EMU area. The single equation nonlinear models are, on the one hand, the nonlinear change point model and, on the other hand, the nonlinear threshold model.

2.2.1 The nonlinear change point model

Given theoretical considerations on the stability of (1), in this subsection we assume that the conditional distribution of Δy_t^k takes the form of the following non-linear regression (structural change point, SC) model:

$$\Delta y_t^k = ax_t + bx_t \mathbf{1}\{\, t > [T\pi]\} + u_t \qquad (2)$$

where $x_t = (1, s_{t-l})'$, $a = (a_0, a_1)$, $b = (b_0, b_1)$. $\mathbf{1}\{.\}$ denotes the indicator function and [a] denotes the greatest integer value smaller than a. The proportion, π, of observations before the break-point is supposed to be constant. Parameter $\pi \in (0,1)$ indexes the relative timing of the structural shift and b indexes the magnitude of the shift. We are interested in testing the hypothesis

$$H_0 \quad : \quad b = 0 \qquad (3)$$
$$H_1 \quad : \quad b \neq 0 \qquad (4)$$

---

[9]The US spread was also included in the regression of this linear model along with the Euro-land spread. However, the US spread appeared to be non-significant at all horizons with no improvement in the R^2 of the regressions or in the standard error. The results are not reported for space considerations but are available from the authors upon request.

This is performed using two parameter stability tests, namely, the supremum Lagrange multiplier ($\sup LM$) statistic of Andrews (1993) and the LM statistic of Andrews and Fair (1988). The $\sup LM$ statistic tests parameter constancy against the alternative of a single unknown breakpoint. Notice that, by construction, the statistic is able to "suggest" where the structural break occurs[10].

The LM statistic assumes that the potential break point is known a priori and it extends classical structural change testing to a wide variety of models. In addition it allows for general forms of temporal dependence and heteroscedasticity in the errors. In our application, the $\sup LM$ statistic takes the form,

$$\sup_{\pi \in \Pi} LM(\pi) = \sup_{\pi \in \Pi} \left(\frac{1}{\pi(1-\pi)} g_1^{'} M^{'} V^{-1} M g_1 \right)$$

where Π is some pre-specified subset[11] of $(0,1)$. g_1 represents orthogonality conditions computed over the first subsample which, in the case of OLS, are $g_1 = (1/T) \sum_{t=1}^{[T\pi]} x_t u_t$ for x_t a $(r \times 1)$ vector of explanatory variables, $r \geq 1$. Moreover, in the OLS case, matrix $M^{'} V^{-1} M$ simplifies to $\hat{S}_T^{-1}$ an estimate of the inverse of the asymptotic variance matrix of $x_t u_t$,

$$S = \lim_{T \to +\infty} (1/T) \sum_{t=1}^{T} \sum_{j=-\infty}^{+\infty} E(u_t u_{t-j} x_t x_{t-j}^{'})$$

Given that equation (1) has moving average errors of order $k-1$ for $k > 1$ the Newey-West (1987) estimate $\hat{S}_T = \hat{\Gamma}_{0,T} + \sum_{j=1}^{q} (1 - \frac{j}{q+1})(\hat{\Gamma}_{j,T} + \hat{\Gamma}_{j,T}^{'})$ of S is employed where $\hat{\Gamma}_{j,T}$ denotes the sample estimate of the jth autocovariance matrix of $x_t u_t$. The truncation lag q is set equal to the optimal value of $[4(T/100)^{2/9}]$ which produced $q = 4$ for $q \leq k$. Due to the presence of moving average errors though, we set $q = k$ when $k > 4$. Critical values for the $\sup LM$ statistic are provided by Andrews (1993).

Preliminary results showed that the $\sup LM$ statistic had difficulties in surpassing even the 10% significance level. The $\sup LM$ statistic rejected the presence of a structural break at all forecasting horizons except for $k = 1, 3$ and 4 where it exhibited marginal significance at the 10% level. A different picture emerged when we examined the LM statistic with break points specified by the date at which the $\sup LM$ statistic was produced. The null hypothesis of parameter constancy was rejected for horizons up to $k = 6$ at least at the 10% significance level. That

---

[10]Our choice of the Andrews (1993) statistic over the Bai (1999) multiple change points statistic was based on the relatively small available sample size of 120 observations and on parsimony reasons. Our aim is to question the stability of the EMU spread-output relationship rather than to pinpoint the exact number of breaks.

[11]We set $\Pi = (\pi_0, 1 - \pi_0)$ with $\pi_0 = 0.20$ so that 20% of the observations from the beginning and end of sample are discarded.

means, as the growth horizon k increases stability of the linear equation (1) cannot be rejected. However, the significance of the spread as a regressor is reducing as k increases and we suspected that the difficulty to identify structural breaks is based on the constancy (or small variation) of the intercept term and not on the slope parameter.

For this reason we calculated the aforementioned statistics under the alternative of partial parameter instability, $H_0 : b_1 = 0 \ vs \ H_1 : b_1 \neq 0$. Numerical results for these $\sup LM$ and LM tests are presented in the last four rows of Table 1. The $\sup LM$ statistic strongly rejects the null hypothesis of stability of the slope parameter at the 5% level for horizons $k = 1$ to $k = 6$. The statistic attains a maximum at the first quarters of 1992 across all k[12]. The same conclusion regarding the stability of the slope parameter is reached when we examine the LM statistic. The null hypothesis of parameter constancy is rejected for horizons up to $k = 6$ at least at the 10% significance level. Overall, there is significant evidence of structural instability in the linear relation (1) that is particularly related to the behavior of the spread coefficient. In the following section we will consider a more general threshold model where not only time but other variables as well are regarded as possible threshold variables.

2.2.2 The nonlinear threshold model

In this section we consider a threshold model which can be viewed as the alternative hypothesis to the linear relationship (1) and it can be specified as,

$$
\begin{aligned}
\Delta y_t^k = {} & (a_0 + a_1 s_{t-l}) \, \mathbf{1}\{z_{t-d} \leq c\} \\
& + (b_0 + b_1 s_{t-l}) \, \mathbf{1}\{ z_{t-d} > c\} + u_t
\end{aligned}
\tag{5}
$$

The idea of approximating a general nonlinear autoregressive structure by a threshold with a small number of regimes was originated by Tong (1983). 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 the spread at time $t - l$.

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

---

[12]The test statistics appeared to be sensitive in the choice of q when calculating the asymptotic covariance matrix of $x_t u_t$. The estimated autocorrelation function of $x_t u_t$ showed there was no significant remaining autocorrelation at lags greater than the chosen q.

since nothing can be learned about c from the data when the null hypothesis is true. The second difficult statistical issue associated with threshold 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. Unfortunately, Hansen's procedures do not apply to models exhibiting correlation[13], a feature which is always present in our case in the form of moving average errors. To the best of our knowledge the complication of serial correlation has not been addressed in the literature so we calculate the marginal significance level of the threshold test statistic by a Monte Carlo method that is described in Appendix 2 below. Peel and Taylor (2002) adopted a similar strategy in a threshold model under the null of random walk.

2.3 Econometric results.

By applying our proposed procedure, we proceed to the specification of model (5). We restrict the delay lag in the set $[0, 1, 2, ..., 12]$ whereas the set of candidate threshold variables included $\{s_t, \omega_{i,t}, \Delta y_t, \Delta_4 y_t, t\}$ with $\omega_{i,t} = \frac{1}{i+1} \sum_{j=0}^{i} s_{t-j}$ and $i = 1, 2, 3$. Endogenous asymmetries in the spread - output relationship are captured by $\{s_t, \omega_{i,t}, \Delta y_t, \Delta_4 y_t\}$ in the sense that certain past spread or growth levels divide the world into two different regimes where the economic significance of spread as an indicator of future growth varies. Following Galbraith and Tkacz (2000) we employed variable $\omega_{i,t}$ in an effort to utilize information based on the course of s_t along a past time interval rather than on the value it admitted on a specific point in time. The same intuition applies with the inclusion of $\Delta_4 y_t$. Exogenous asymmetries in the spread - output relationship are investigated using time t as a possible threshold variable. Delay lag in this case is always set equal to zero and model (5) is interpreted as a change point model[14]. Table 2 presents estimates of the delay lag that minimizes the residual sum of squares while Table 3 presents the F statistic for the null hypothesis of no threshold behavior.

Our results suggest: a) there is evidence of significant nonlinearity (structural break) with respect to time that seems to disappear when the forecasting horizon exceeds 6 quarters or $1\frac{1}{2}$ years ahead in accordance with the $\sup LM$ and LM tests of the previous section, b) there is evidence of asymmetry with respect to past growth levels, in particular with respect to annual growth at lags of 9 quarters or

---

[13]We thank an anonymous referee for pointing out this fact.

[14]The search interval for transition variable z_{t-d} was set to be $C = [\mathfrak{c}, \bar{c}] = [\min\{z_{t-d}\} + 0.1, \max\{z_{t-d}\} - 0.1]$ whereas for the case $z_t = t$ we set $C = [\mathfrak{c}, \bar{c}] = [\min\{z_t\} + 10, \max\{z_t\} - 10]$. Notice that in order to "gain" observations, we arbitrary add or subtract 0.1 (or 10) to construct the boundaries $\mathfrak{c}, \bar{c}$ and then we divide C into 100 discrete points. By following the aforementioned strategy, we did not experience difficulties in the solution of problem (13) (see Appendix 2). An alternative approach would be to discard a selected percentage (usually 10% or 15%) of the observations in both ends of z_{t-d}.

around 2 years in the past, albeit of lower significance than with respect to time[15] c) there is no evidence of asymmetries with respect to past spread levels at five percent level. The robustness of this result was further checked by splitting the sample at the suggested break dates and testing for asymmetries with respect to the spread before and after the sample split. We were still unable to find significant nonlinearities.

Consequently, we set $z_t = t$ (time threshold or change point model) and $z_{t-d} = \Delta_4 y_{t-d}$ (growth threshold model) and we proceed with parameter estimation in model (5). The results are tabulated in Tables 4 and 5 respectively. Both tables suggest improved fit of the nonlinear model judging by the reduction in the equations' standard error $s.e.$. For horizon $k = 6$ the nonlinear models achieve a reduction up to 20% (the linear and nonlinear models standard error ratio is close to 0.80).

Regarding Table 4, the estimated breakpoints coincide for horizons $k \geq 2$ and they are located at the end of 1992. The dynamics implied by our estimates suggest that the spread has significant positive effects on real activity only before the end of 1992 (around 70% of our sample size). The slope coefficient magnitude varies from a minimum of 1.15 at the $k = 6$ quarter horizon to a maximum of 1.87 at $k = 1$. After 1992, the significance of the multiplier decreases to zero both in statistic and economic terms. This finding is consistent with a recent theoretical explanation regarding the real activity - spread relationship put forward by Estrella (1998). The relationship is weak or nonexistent when the central bank targets inflation only. Such target practices have been adopted in EU countries mostly during the 1990's.

Regarding Table 5, our results point to asymmetry around relatively large levels of past annual growth. The threshold estimates are and $\hat{c} = 3.38$ for $k = 2, 3, 4, 5, 6$ (63 obs of $\Delta_4 y_{t-9}$ are below that threshold and 55 above). Even if other thresholds regarding past levels of growth exist, the positive bound around 3.38% of annual growth two years ago signifies the most important threshold in terms of reducing the residual sum of squares of the forecasting equation. Above that levels of growth, spread enters significantly in the equation with minimum slope parameter estimate $\hat{b}_1 = 0.97$ at $k = 6$ and maximum slope estimate $\hat{b}_1 = 1.74$ at $k = 2$. Usually, a growth rate above 3.38% implies a business cycle peak (see Figure 1a). Over two years after a business peak (k=9), the economy is usually at a low growth pace and monetary authorities could be more concerned with output growth rather than inflation. In such situations the spread has the ability to predict future output growth. This result is consistent with Estrella (1998) argument explained above. The estimated threshold of 3.38% p.a. seems to designate, on average, the point that, when crossed, the Central Bank's attention focuses to output movements[16].

---

[15]For $k = 1$ - with s_t or $\omega_{1,t}$ being the threshold variables - our results do not differ significantly from those of Galbraith and Tkacz (2000). However, Galbraith and Tkacz did not consider time or output growth as candidate threshold variables.

[16]Notice that the results tabulated in tables 4 and 5 regarding multiplier estimates $\hat{a}_1, \hat{b}_1$ are not in contradiction. As a further aid to interpretation of these results, figure 2 depicts an informative scatter diagram of annual growth 4 quarters into the future Δy_t^4 against the spread (s_{t-1}) for the

2.4 Comparing forecasting accuracy

In this section we conduct a forecasting experiment. Our purpose is to compare the out-of-sample forecasts of the two alternative models (1) and (5). Several studies have compared the forecasting ability of linear versus nonlinear TAR models (see e.g. Aydemir, 1998; and Boero and Marrocu, 2001). In most of the cases, the nonlinear models do not outperform the linear ones. It is interesting to see whether such conclusions can also be drawn for the spread - output relationship. We retain $m = 64$ observations for out-of-sample forecasting. In particular we will calculate one step ahead forecasts for the period 1986.01 onwards. In order to reduce parameter uncertainty and adopt a policy making view, both linear and nonlinear models were re-estimated for each observation added to the sample after the date 1986.01 and the forecasts are based on the "updated" parameter estimates. Thus, for each quarter after 1985:4, we add one observation to the sample, a searching algorithm is applied across candidate threshold values and when the residual sum of squares has been minimized, the corresponding threshold model is estimated and one period ahead forecasts are produced.

The forecast evaluation is based on two well known different criteria, namely, the mean square prediction error (MSPE) and the mean absolute prediction error (MAPE). Although these criteria 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 statistically significant results on the difference of competing forecasts. In order to test the null hypothesis of equal forecasting accuracy we will employ two different tests developed by Diebold and Mariano (1995). 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 (loss) produced by two different models. For our application we chose the quadratic function,

$$d_j = \left(\hat{y}_{T_0+j|T_0+j-1,NL} - y_{T_0+j}\right)^2 - \left(\hat{y}_{T_0+j|T_0+j-1,L} - y_{T_0+j}\right)^2 \qquad (6)$$

although the absolute difference function did not qualitatively alter our results. Subscript indices NL and L denote nonlinear and linear forecasts respectively.

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

$$S = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \left(\mathbf{1}\{d_j > 0\} - 0.5\right) \sim N(0,1) \qquad (7)$$

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. Non-rejection of the null hypothesis would imply a zero-median loss differential or else if $e_{NL,j} = \left(\hat{y}_{T_0+j|T_0+j-1,NL} - y_{T_0+j}\right)^2$ and $e_{L,j} = \left(\hat{y}_{T_0+j|T_0+j-1,NL} - y_{T_0+j}\right)^2$ we have $P(e_{NL,j} > e_{L,j}) = P(e_{NL,j} < e_{L,j})$.

---

estimation sample period $t \leq 1992Q4$, $\Delta_4 y_{t-9} > 3.38$. It is worth noticing that figure 2 shows a strong visual positive relationship between spread and output growth.

Diebold and Mariano (1995) propose the use of an asymptotic statistic which tests whether the average loss differential is significantly different from zero and allows for serially correlated loss differentials. Specifically they show that,

$$DM = \frac{\bar{d}}{\sqrt{\hat{\omega}/m}} \rightarrow_d N(0,1) \tag{8}$$

where $\hat{\omega}$ is an estimate of the long run covariance matrix of d_j[17]. The forecasting period is 1986:1 - 2000:4 and given the threshold variables, the initial sample period employed to commence the one step ahead forecasting experiment was 1970:1 - 1985:4 for the linear and time threshold models and 1973:3 - 1985:4 for the growth threshold model. Table 6 reports out-of-sample forecast evaluation results. As indicated by the MSPE and MAPE criteria, the loss function chosen does not impact on the forecast evaluation. Based on those criteria both nonlinear models perform better than the standard linear one in all growth horizons and the structural change model seems to perform better at short to mid - term growth forecast horizons. This is overturned by the threshold model for $k = 5,6$. Following the S and DM forecast equality statistics, the latter model performs better than the structural change model since we can reject the null hypothesis for forecast equality for horizons $k = 1,2,3,5,6$ whereas the SC model rejects only at $k = 2,3$.

Most of this success is due to the "overshooting" during the post-1992 period of the linear model forecasts. The linear model still assumes large spread multipliers although the significance of the spread according to both time and growth threshold models has been dramatically reduced. As a further visual aid, Figures 3 and 4 depict actual, nonlinear and linear forecast values for the time and growth threshold models respectively at $k = 4$. We observe that the time threshold model in the post-1992 period simply updates the mean and ignores deviations in actual growth movement. On the other hand, the growth threshold model also spends the majority of the forecasting time interval updating the mean but there were circumstances when the upper regime parameters were activated due to past annual growth exceeding the estimated threshold of 3.38%. Overall, the nonlinear models outperform the linear one in terms of point forecasting accuracy. Next section examines a different question. The predictive ability of the yield spread over recession periods instead of point estimates.

---

[17]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 $\hat{\omega} = \hat{\gamma}_0(d)$, (the variance of d_j), when the DM test is applied to 1-step ahead forecasts. Nevertheless, our model is nonlinear and although $h = 1$ we found that d_j was persistent with significant correlations up to a certain order (in almost all cases up to $\hat{\gamma}_3$). For example, comparing the linear and threshold models when $k = 4$ (and the threshold is $\Delta_4 y_{t-9}$), $\hat{\gamma}_1(d) = 0.764$, $\hat{\gamma}_2(d) = 0.447$ and $\hat{\gamma}_3(d) = 0.221$ (fourth order correlation was $\hat{\gamma}_4(d) = 0.10$). 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.

3 The probability of recessions

Estrella and Hardouvelis (1991), and Estrella and Mishkin (1997, 1998), applied probit models to show the usefulness of the yield curve as an indicator of future recession. The conclusion they arrived at was that the yield curve dominates other leading indicators and a number of macroeconomic time series. Camacho and Perez-Quiros (2002) found that a combination of switching regime and nonparametric models outperform a number of alternatives, including the probit model, when predicting both the NBER business cycle and GDP growth for the US economy[18].

In the European context, Bernard and Gerlach (1996), applying the probit methodology to eight industrialized countries[19], conclude that the yield curve provides information about the likelihood of a future recession in all eight countries. Obviously, there are notable differences between countries, perhaps due to differences in financial markets regulations. More recently, Sensier et al. (2002) have documented that financial and real variables predict recessions in major European countries using logit models. Moreover, they show the influence of international events to predict European business cycle regimes. In particular, the US leading indicator index and short rate interest rates. In this section we shall concentrate on details concerning the specification of a probit model using the Euro area aggregate data and subsequently on the empirical results.

3.1 Probit specification

We adopt the following specification to predict a binary variable X_t that takes a unit value when there is a recession and zero otherwise:

$$\Pr[X_t = 1 | z'_{t-k}] = F(\beta' z_{t-k}) \tag{9}$$

where Pr denotes probability, $X_t = 1$, is an indicator variable of recessions, and F is the cumulative normal distribution function that takes real values and returns values ranging from zero to one. This specification belongs to a standard probit model, which is non-linear and relates the indicator variable X_t to a vector of exogenous variables z; in our case, a constant and the EMU yield curve. Sensier et al. (2002) have suggested that the US yield spread also has predictive ability over European business cycle turning points. To check this suggestion on aggregated data for the EMU we will also include the US spread in the exogenous vector z_t. The log-likelihood function of this model is:

---

[18]An alternative way of extracting information on predictive probabilities in a multi-equation context is proposed by Anderson and Vahid (2000). However, we do not apply this methodology due to the fact that our nonlinear models are not autoregressive.

[19]These countries are Belgium, Canada, France, Germany, Japan, Netherlands, UK and US.

$$L = \sum_{t=1}^{T} \left\{ X_t \ln F(\beta' z_{t-k}) + (1 - X_t) \ln(1 - F(\beta' z_{t-k})) \right\}$$

where the integer k measures the time lag. and the fitted values $F(\hat{\beta}' z_{t-k})$ can be interpreted as the probability that a recession will occur, conditional on the observed value of the spread.

Dating recessions. Given that NBER type of turning points identifying classical business cycles is not available for our series we rely on a measure of recession reflecting growth cycles and calculated according to the following simple rule (Fair, 1993): quarter t is in recession when it has negative growth and there are at least two consecutive quarters of negative growth in the window from t to $t+4$. The rule produced the following recession dates: 1974:4 - 1975:1, 1980:4 - 1981:1, 1982:3 - 1982:4 and 1992:3 - 1993:2[20].

The main shortcoming of model (9), as mentioned by Dueker (1997), is the fact that the traditional probit estimation can be mis-specified if there is information content in the autocorrelation structure of the binary time series. It is implausible to assume that reference to whether the economy has actually been in recession in recent periods is non-important. Therefore, we expand (9) and we also estimate

$$\Pr[X_t = 1 | z'_{t-k}, X_{t-k}] = F(\beta' z_{t-k} + \gamma X_{t-k}) \tag{10}$$

which is able to test for information content in the autoregressive structure of the binary time series at the moment when the indicator gave a signal. The choice of k is based on the modified McFadden's pseudo R^2, (PR^2) that measures the explanatory power of the model[21]. The best estimates were obtained with three lags on the EMU spread, twelve on the US spread and four lags on the lagged dependent variable.

3.2 Empirical results

Table 7 shows the estimation output for equations 9 and 10. Overall the estimation output, as is shown in Table 7, suggests the ability of the Euro area and US yield curves to predict recessions in EMU three quarters ahead. The coefficients of lagged spread are negative as expected and statistically significant. A negative sign of the spread estimated coefficients means that an increase in the term spread reduces the probability of a future recession whereas the sign of $\hat{\gamma}$ negatively relates the future probability of a recession to the state of the economy one year ago. If

---

[20]Alternative definitions of recessions in the Euro area with similar results can be found in Ross and Ubide (2001) and Krolzig and Toro (2001). For an application of probit models using these alternative recession dates see Moneta (2003).

[21]For example, in the case of model (10), 1728 (=12 × 12 × 12) different specifications were estimated varying k from 1 to 12 for each variable of the model.

the economy was in recession a year ago is more likely that in future quarters the economy will not be in recession[22].

Figures 5-7 illustrate the estimation results shown in Table 7. Figure 5 plots the actual term spread against the recession dates represented by the shaded areas. The figure shows the role of the spread as leading indicator for recessions as the spread becomes negative before all shaded areas. Figures 6 and 7 plot the fitted values from estimation of models (9) and (10); in other words, this is the within sample prediction of a future recession. The cross hatched areas denote periods of actual recessions as defined above. As the figures illustrate, all peaks in the estimated probability were associated with a recession. The models predict three quarters ahead major recessions during the sample period. In particular, model (9) predicts the recessions in 1974:4-1975:1, 1980:4-1981:1 and 1982:3 - 1982:4 with a probability of more than 50%. The model predicts with more that 80% the collapse of the European Monetary System (EMS) and subsequently the recession of 1992. Model (10) produces higher probability estimates - around 70% - in all recession dates and improves the pseudo R^2.

We also examine the "out-of-sample" prediction of future recessions and check whether the recession of 1992, when the EMS was disintegrated, could be predicted. This experiment focuses on the merits of such simple binary models as recession predictors. This information is illustrated in Figures 8-9. Thus, while Figures 6 and 7 plot the fitted values from probit estimation for the period 1970-2000, Figures 8 and 9 plot the fitted values of each model based on one-step recursive estimation from 1991:1 up to 2000:4. Model (9) produces a probability of recession of 51.52% at the third quarter of 1992 which is identified as the first quarter of the last recession period. Probability increases gradually up to 88.96% in 1993:2 which is actually the last quarter of that recession period. The probability of recession remains above 50% throughout 1993 and then sharply decreases to zero. On the other hand, model (10) forecasts the probability of recession in the third quarter of 1992 to be 57.27%, reaches its maximum of 91.22% at the last quarter of the recession period and then drops immediately to zero.

4   Conclusions

Our general conclusion is that the term spread of the Euroland is a useful indicator of future output growth and of future recessions in the Euro area. Linear indicator and non-linear threshold indicator models predict reasonably well output growth four quarters ahead which has been taken as the "stylized" real growth horizon. However, the linear model shows signs of instability over the sample period covered by our dataset, 1970:1 - 2000:4. On the grounds of these results we have estimated two alternative nonlinear models. The first one is a nonlinear change point model where we find a significant structural break in 1992. The second one is a threshold model which includes asymmetric effects in terms of lagged annual

---

[22]We shall recall that within our sample period recessions do not last longer than four quarters.

growth rate and find that at times of slow growth, once the economic 'boom' has died out, the yield spread is a more effective leading indicator. With regard to recursive out-of-sample point forecasts, the nonlinear models produce more accurate values of future output growth in the Euro area, either in terms of MSPE's and MAPE's forecast evaluation criteria or Diebold and Mariano (1995) tests.

We have also confirmed the power of the Euroland and US yield curves in predicting recessions in Europe. This result is supported by both in-sample and "out-of-sample" probit predictions. In both cases, the model predicts at least three quarters ahead major recessions in the Euro area in the last thirty years.

Appendix 1

The creation of the EMU has also created the necessity to obtain aggregate statistics for this economic area. Up to now, as far as we know, official bodies have not published a homogeneous database with appropriate frequencies and covering a period long enough to carry out statistical and econometric studies. However, the European Office of Statistics, Eurostat, and the European Central Bank, have just published official data for the EMU starting from the mid 1990's and late1998, respectively. For this reason, in the present work we have decided to use the database that was kindly facilitated by the Research Department of the Banco Bilbao Vizcaya Argentaria (BBVA)[23]. This database includes the initial eleven countries: Austria, Belgium, France, Germany, Holland, Italy, Ireland, Luxemburg, Portugal, Finland, and Spain. Also, there are two major reasons for using this database. One is that it is based on OECD, Main Economic Indicators and Eurostat. The second reason is the methodological advantage of a more flexible method of aggregation.

This method is flexible because its weights take into account the specific nature of the variable under study, that is whether it is real, nominal or an index number[24]. In our particular case, we use quarterly data over the period 1970:1 - 2000:4 for real GDP seasonally adjusted, 10 years government bond yield, and the 3-month interest rate on deposits. Real GDP has been aggregated converting each country specific GDP into a common currency[25]. The weights used are PPP based for 1990. Also, to preserve the dynamics of the original series and avoid data contamination as a result of exchange rate variability a fixed conversion from the year 1990, as a base year, has been chosen. In the case of interest rates, the weights are simply each country GDP over total GDP.

---

[23]Further details on database can be found in Ballabriga and Castillo (2003).

[24]For a detailed explanation on the advantages of PPP based GDP measures see Gulde and Shulze-Ghattas (1992).

[25]Alternative weighting schemes can be found in Fagan and Henry (1998).

Appendix 2

We can write the threshold model (5) compactly as,

$$\Delta y_t^k = x_t(c)^{'} \theta + u_t \tag{11}$$

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

$$\hat{\theta}(c) = \left(\sum x_t(c) x_t(c)^{'} \right)^{-1} \left(\sum x_t(c) \Delta y_t^k \right) \tag{12}$$

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,

$$\hat{c} = \arg \min_{c \in C} \sigma_T^2(c) \tag{13}$$

where C is an interval (usually trimmed) that covers the sample range of the transition variable. The specification and testing of (5) proceeds in three steps.

Step 1: Problem (13) is solved by a direct search over C and the LS estimate of θ is denoted, $\hat{\theta} = \hat{\theta}(\hat{c})$. The LS principle allows us to estimate the, typically, unknown value of d by extending problem (13) to a search across the discrete space $[1, \bar{d}]$.

Step 2: For the choice of d that minimizes the residual variance, the hypothesis $H_0 : \alpha = \mathbf{b}$ is tested as follows: Let $\{\eta_t\}_{t=1}^{T}$ be an $i.i.d$ sequence of $N(0,1)$ draws. Let V denote the Toeplitz covariance matrix of $\hat{u}_t$ under H_0. Using the Choleski decomposition of $V = RR^{'}$ create $e_t = R\eta_t$ that has the same covariance structure as $\hat{u}_t$. 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). The method will not provide confidence intervals for $\hat{c}$.
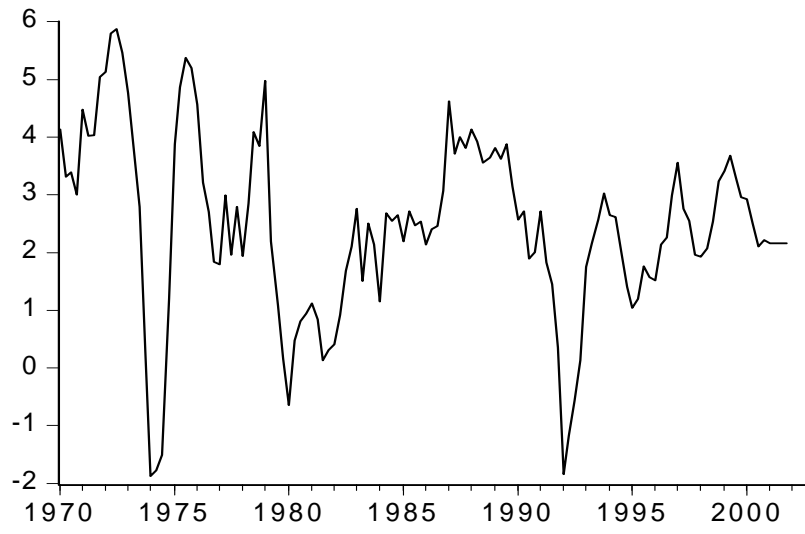
18

References

[1] Anderson, H.M., & Vahid, F. (2000). Predicting the probability of a recession with nonlinear autoregressive leading indicator models. Monash University Working Paper 3/2000

[2] Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, **61**, (4), 821-856

[3] Andrews, D.W.K., & Fair, R.C. (1988). Inference in nonlinear econometric models with structural change. *The Review of Economic Studies*, **55**, (4), 615-639

[4] Aydermir, A.B. (1998). Forecast performance of threshold autoregressive models- A Monte Carlo study. *UWO Department of Economics Working Papers* 9910.

[5] Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics*, **91**, 299-323.

[6] Ballabriga, F.C., & Castillo, S. (2003). BBVA-ARIES: A Forecasting and Simulation Model for EMU. *Journal of Forecasting*, **22**, 411-426.

[7] Berk, J.M., & Van Bergeijk, P. (2000). Is the yield curve a useful information variable for the Eurosystem. *European Central Bank. Working paper No*. 11.

[8] Bernard, H., & Gerlach, S. (1996). Does the term structure predict recessions? The international evidence. *BIS Working Paper, No.* 37, BIS, Basle.

[9] Blanchard, O. J., & Fischer, S. (1989). Lectures on Macroeconomics. The MIT Press, Cambridge (MA).

[10] Boero, G., & Marrocu, E. (2001). Evaluating non-linear models on point and interval forecasts: an application with exchange rate returns. *CRENOS Working Paper* 01/10.

[11] Camacho, M., & Perez-Quiros, G. (2002). This is what the leading indicators lead, Journal of Applied Econometrics, **17**, 61-80.

[12] Davis, E.P., & Fagan, G. (1997). Are financial spreads useful indicators of future inflation and output growth in E.U countries. Journal of Applied Econometrics, **12**, 701-714.

[13] Diebold, F.X., & Mariano, S. (1995). Comparing predictive accuracy. *Journal of Business and Economics Statistics,* **13**, 253-263.

[14] Dotsey, M. (1998). The predictive content of the interest rate term spread for future economic growth. Federal Reserve Bank of Richmond *Economic Quarterly*, Vol. 84(3), 31-51

[15] Dueker, M. (1997), Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions, Federal Reserve Bank of St. Louis Review, No. 2, 41-51.

[16] Estrella, A. (1998). Monetary policy and the predictive power of the term structure of interest rates. Federal Reserve Bank of New York, November

[17] Estrella, A., & Hardouvelis, G. (1991). The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance*, **46,** (2), 555-576.

[18] Estrella, A., & Mishkin, F. S. (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

[19] Estrella, A., & Mishkin, F. S. (1998). Predicting US Recessions: Financial Variables as Leading Indicators. *The Review of Economics and Statistics*, **80**, 45-61.

[20] Estrella, A., Rodrigues, A.P., & Schich, S. (2000). How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States. *Mimeo, Federal Reserve Bank of New York.*

[21] European Central Bank (1999), Monthly Bulletin, July 1999, European Central Bank, Frankfurt am Main, p.21.

[22] Fagan, G., & Henry, J. (1998), Long-run money demand in the E.U: Evidence for area-wide aggregates, Empirical Economics, **23**, 483-506.

[23] Fair, R. C. (1993). Estimating event probabilities from macroeconometric models using stochastic simulation, in J. H. Stock and M. W. Watson (eds), Business Cycles, Indicators, and Forecasting, NBER, University of Chicago Press, Chicago, pp. 157-178.

[24] Friedman, B.M., & Kuttner, K.N. (1998). Indicator Properties of the Paper-Bill Spread: Lessons from Recent Experience. *The Review of Economics and Statistics*, **80**, 34-44.

[25] Fuhrer, J.C., & Moore, G. R. (1995). Monetary Policy Trade-offs and the Correlation between Nominal Interest Rates and Real Output. *American Economic Review*, **85**, 219-239.

[26] Galbraith, J.W., & 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.

[27] Galvao, A. (2003). Structural Break Threshold VARs for Predicting US Recessions using the Spread. *Ibmec Working Paper* - WPE - 2003 - 4
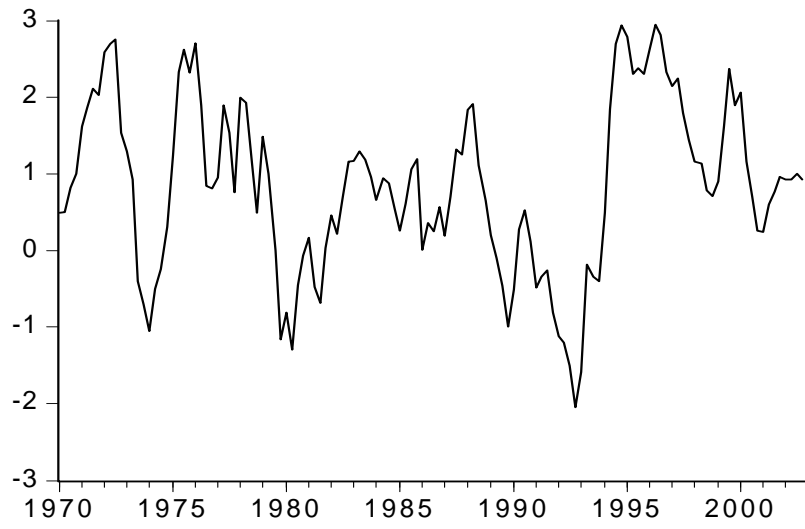
[28] Gulde, A.M., & Shulze-Ghattas, M. (1992). Aggregation of economic indicators across countries: Exchange rate versus PPP based GDP weights. IMF Working Paper no. 36.

[29] Hamilton, J.D., & Kim, D.H. (2002). A re-examination of the Predictability of Economic Activity Using the Yield Curve Spread. *Journal of Money, Credit & Banking*, **34**, 340-360

[30] Hansen, B.E. (1997). Inference in TAR Models. *Studies in Nonlinear Dynamics and Econometrics, **2**, 1-14.*

[31] Haubrich, J.G., & Dombrosky, A. M. (1996). Predicting Real Growth Using the Yield Curve, *Federal Reserve Bank of Cleveland Economic Review*, **32** (1) 26-34.

[32] Krolzig, H.M., & Toro, J. (2001). Classical and modern business cycle measurements: The European case. Institute of Economics and Statistics Oxford Discussion Paper No. 60.

[33] Modigliani, F., & Sutch, R. (1966). Innovation in interest rate policy. *American Economic Review, Papers and Proceeding, 76, 178-197.*

[34] Moneta, F. (2003). Does the yield spread predict recessions in the euro area?. ECB Working Paper No 294.

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

[36] Peel, D.A., & Taylor, M.P. (2002). Covered interest rate arbitrage in the interwar period and the Keynes-Einzig conjecture, *Journal of Money, Credit & Banking*, 51-85

[37] Plosser, C.I., & Rouwenhorst, K.G. (1994). International Term Structure and Real Economic Growth. *Journal of Monetary Economics, **33,** 133-155.*

[38] Ross, K., & Ubide, A. (2001). Mind the gap: What is the best measure of slack in the Euro area?. IMF Working Paper December 2001.

[39] Sensier, M., M.J. Artis, D.R. Osborn and C.R. Birchenhall (2002). Domestic and international influences on business cycle regimes in Europe. Centre for Growth and Business Cycle Research, Discussion Paper 11, University of Manchester, *forthcoming International Journal of Forecasting*.

[40] Stock, J.H., & Watson, M.W. (1989). New Indices of Coincident and Leading Indicators, In O. Blanchard and S. Fisher edited NBER Macroeconomic Annual**.** Cambridge, MIT Press.

[41] Stock, J. H., & Watson, M.W. (2001). Forecasting output and inflation: The role of asset prices, NBER Working Paper 8180.

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

[43] Tong, H. (1983) Threshold Models in Non-linear Time Series Analysis, Springer-Verlag: New York.

[44] Turnovsky, S. J. (1989). The term structure of interest rates and the effects of Macroeconomic Policy. *Journal of Money, Credit and Banking, 21, (3), 321-347.*

[45] Venetis, I.A., Paya, I., & Peel, D.A. (2003). Re-examination of the predictability of economic activity using the yield spread: A nonlinear approach. *International Review of Economics and Finance, 12, (2), 187-207*

(a) Growth rate of real GDP

Figure 1(a)

(b) Interest rate spread

Figure 1(b)

23

Figure 2

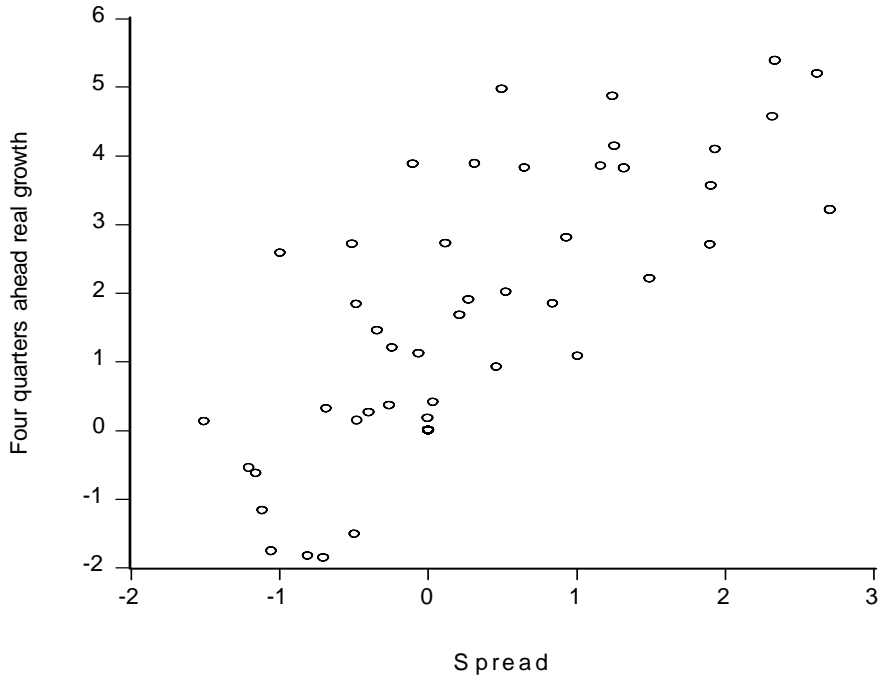Period 1986:1 - 2002:4. Time threshold model (SC)

Figure 3

24

Forecastig Period: 1986:1 - 2002:4. Growth threshold model
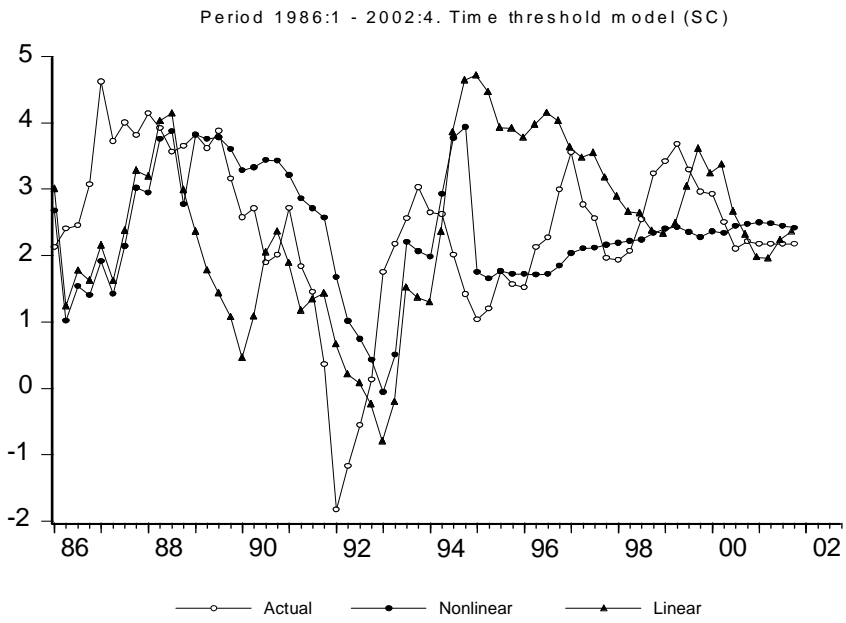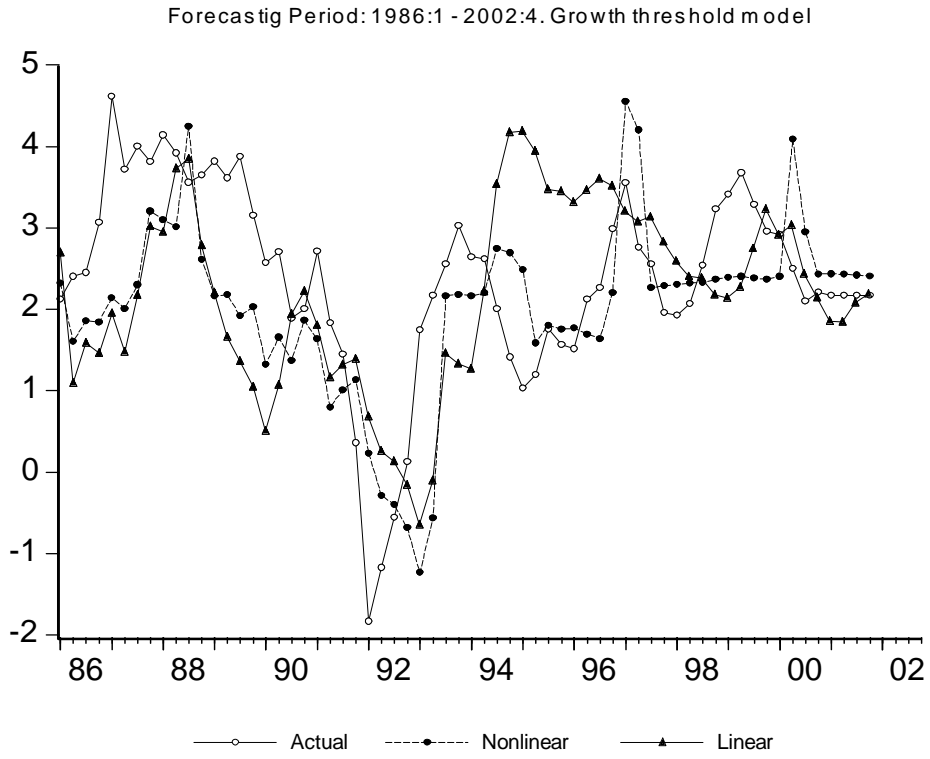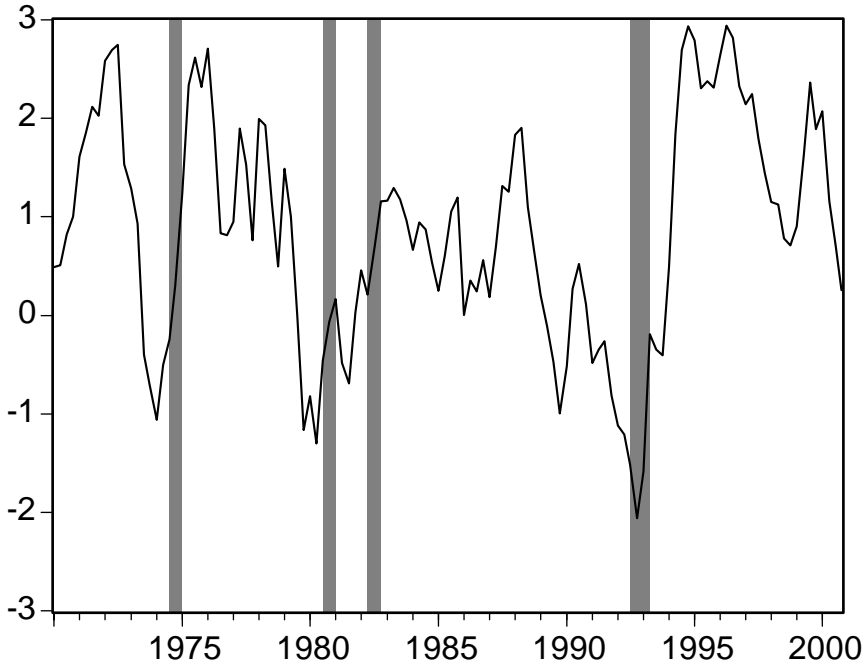

Figure 4

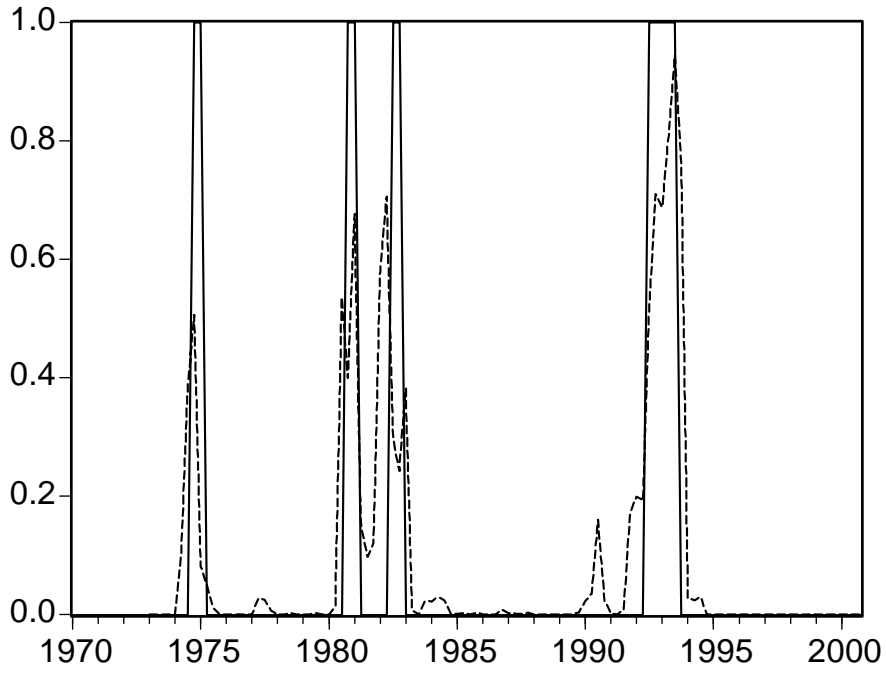Figure 5. Spread and recession dates

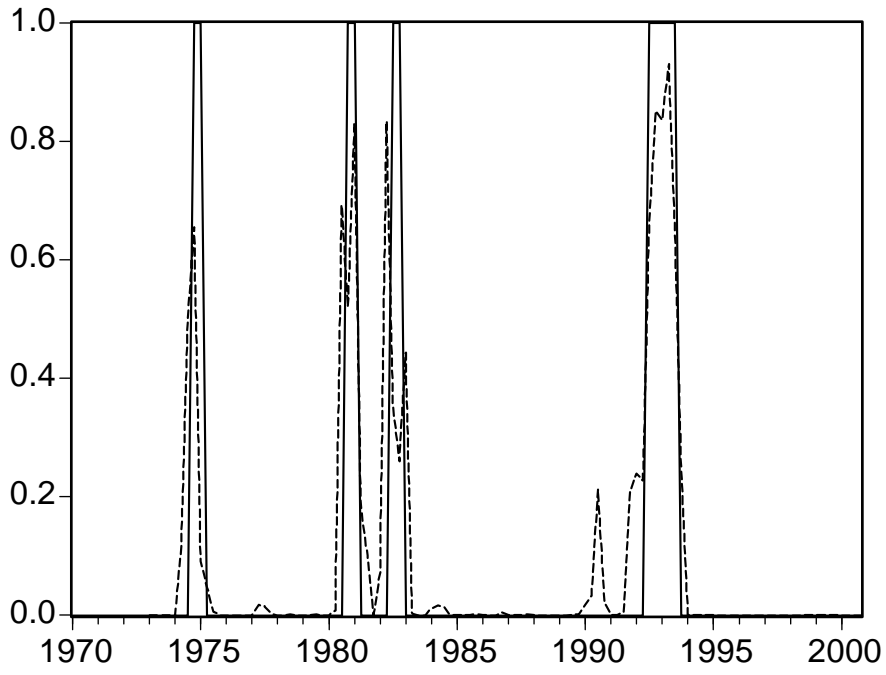Figure 6. Fitted values of model (9)
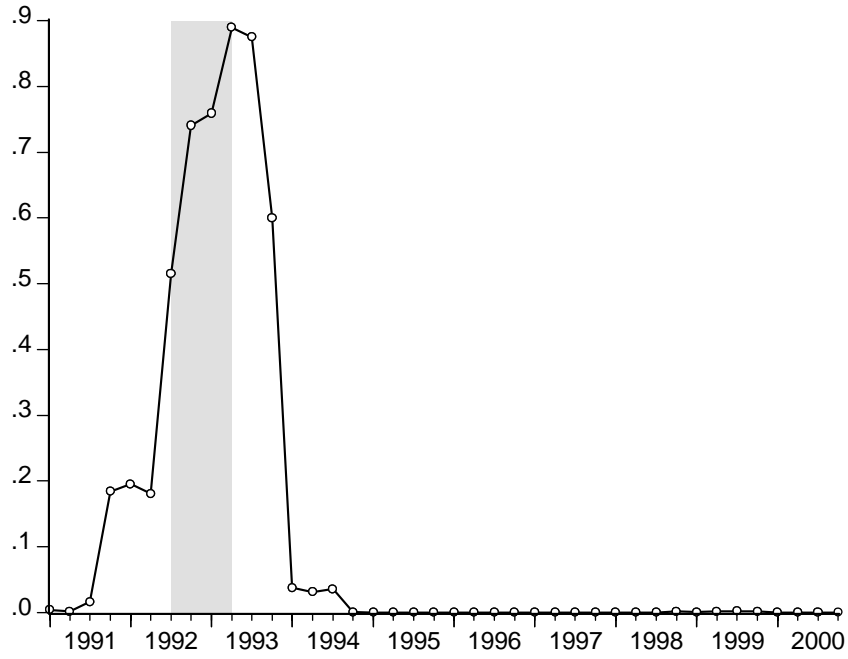

Figure 7. Fitted values of model (10)

Figure 8. One-step forecasts of the probability of recession using model (9).
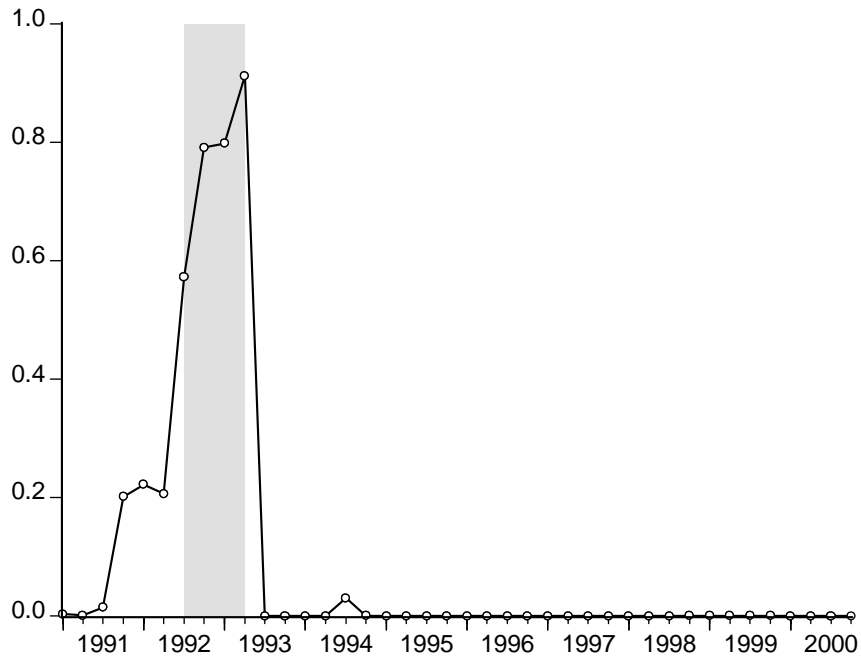Forecasting period: 1991:1 2000:4

Figure 9. One-step forecasts of the probability of recession using model (10).
Forecasting period: 1991:1 2000:4

27

Table 1. OLS estimates of $\Delta y_t^k = a_0 + a_1 s_{t-l} + u_t$.

k	1	2	3	4	5	6	8	12
l	2	1	1	1	1	1	1	1
$\hat{\alpha}_1$	0.875	0.863	0.806	0.732	0.671	0.600	0.448	0.325
	(0.268)	(0.257)	(0.244)	(0.227)	(0.206)	(0.183)	(0.146)	(0.096)
R^2	0.156	0.243	0.282	0.278	0.271	0.249	0.187	0.154
s.e	2.297	1.718	1.458	1.324	1.241	1.178	1.054	0.882
supLM	9.941*	9.412*	10.150*	10.457*	9.448*	8.844*	7.506	5.189
Date	1992:3	1992:2	1992:2	1992:2	1992:2	1992:2	1992:1	1992:2
LM	8.536	6.831	7.382	6.877	5.784	4.826	3.118	0.782
p-value	[0.014]	[0.032]	[0.024]	[0.032]	[0.055]	[0.089]	[0.210]	[0.676]

Notes: k denotes quarters ahead growth and l the lag of the spread employed. Numbers in parentheses are Newey - West standard error estimates. s.e denotes the residual standard error. The sup LM statistic of Andrews (1993) was calculated using $\pi_0 = 0.20$ where π_0 indicates that the first subsample is of size $T\pi_0$ and the last subsample is of size $T(1-\pi_0)$. The dimension of the parameter vector tested for stability is p=1 (the slope parameter), and the 5% asymptotic critical value from Andrews (1993, p. 840, TABLE I) is: 8.45. Consequently, one asterisk indicates statistical significance at the 5% level. The row named "Date" gives the sample date at which a) the sup LM statistic was obtained b) the LM statistic was calculated.

Table 2. Searching for the appropriate lag in candidate threshold variables

k		1	2	3	4	5	6	8	12
s_{t-d}	$d:$	9	9	9	9	9	9	9	7
$\omega_{1,t-d}$	$d:$	9	9	9	9	9	9	9	9
$\omega_{2,t-d}$	$d:$	9	9	9	9	9	9	9	9
$\omega_{3,t-d}$	$d:$	9	9	9	9	9	9	9	8
Δy_{t-d}	$d:$	10	10	10	7	10	10	10	9
$\Delta_4 y_{t-d}$	$d:$	9	9	9	9	9	9	9	8
$t-d$	$d:$	0	0	0	0	0	0	0	0

Notes: the table reports the delay lag d which produced the minimum (across d) value in problem (13)

Table 3. Testing the null hypothesis of no threshold behavior. F - statistic

k	1	2	3	4	5	6	8	12
s_{t-d}	**[0.049]**	[0.268]	[0.259]	[0.217]	[0.240]	[0.172]	[0.311]	[0.880]
$\omega_{1,t-d}$	**[0.096]**	[0.269]	[0.198]	**[0.090]**	**[0.060]**	**[0.068]**	[0.281]	[0.794]
$\omega_{2,t-d}$	[0.190]	[0.339]	[0.211]	[0.111]	[0.104]	[0.160]	[0.412]	[0.773]
$\omega_{3,t-d}$	[0.217]	[0.323]	[0.213]	[0.093]	[0.101]	[0.145]	[0.400]	[0.834]
Δy_{t-d}	[0.547]	[0.205]	[0.075]	[0.194]	**[0.081]**	[0.157]	[0.434]	[0.632]
$\Delta_4 y_{t-d}$	**[0.018]**	**[0.024]**	**[0.020]**	**[0.011]**	**[0.021]**	**[0.078]**	[0.275]	[0.235]
t	**[0.008]**	**[0.014]**	**[0.001]**	**[0.012]**	**[0.017]**	**[0.023]**	[0.107]	[0.154]

Notes: Numbers in squared brackets are p-values for the F statistic testing the null hypothesis of no threshold. The lag d of the threshold variables was chosen according to table 2. Bolded types denote statistical significance at the 10% or lower significance level.

Table 4. Threshold model (2). Threshold variable: t. Dependent variable Δy_t^k

k :	1	2	3	4	5	6
$\hat{a}_0$	0.92	1.54	1.55	1.69	1.69	1.73
	(0.37)	(0.18)	(0.15)	(0.12)	(0.12)	(0.11)
$\hat{a}_1$	1.87	1.47	1.41	1.27	1.23	1.15
	(0.26)	(0.15)	(0.12)	(0.10)	(0.10)	(0.09)
$\hat{b}_0$	2.16	2.26	2.39	2.48	2.30	2.32
	(0.30)	(0.36)	(0.29)	(0.29)	(0.24)	(0.22)
$\hat{b}_1$	0.21	0.03	-0.04	-0.08	0.01	0.00
	(0.21)	(0.20)	(0.16)	(0.16)	(0.13)	(0.12)
$\hat{c}$	56	92	92	92	91	91
Date:	1983:4	1992:4	1992:4	1992:4	1992:3	1992:3
$s.e$	2.12	1.51	1.20	1.06	0.98	0.91

Notes: Numbers in parentheses report standard errors. Equation standard error is reported in row "$s.e$". The row named "Date" contains the break dates according to the threshold estimate $\hat{c}$.

Table 5. Threshold model (2). Threshold variable: $\Delta_4 y_{t-d}$. Dependent variable Δy_t^k

k :	1	2	3	4	5	6
$\hat{a}_0$	2.20	2.09	2.06	2.42	2.43	2.42
	(0.30)	(0.21)	(0.17)	(0.18)	(0.17)	(0.16)
$\hat{a}_1$	0.28	0.29	0.27	0.00	0.01	0.03
	(0.21)	(0.15)	(0.12)	(0.13)	(0.12)	(0.11)
$\hat{b}_0$	0.93	1.05	1.21	1.47	1.50	1.54
	(0.41)	(0.31)	(0.25)	(0.15)	(0.14)	(0.13)
$\hat{b}_1$	1.81	1.74	1.59	1.24	1.04	0.97
	(0.33)	(0.23)	(0.19)	(0.12)	(0.11)	(0.11)
$\hat{c}$	4.73	3.38	3.29	3.38	3.38	3.38
$s.e$	2.12	1.52	1.23	1.05	0.98	0.91

Notes: Numbers in parentheses report standard errors. $s.e$ denotes equation standard error. $\hat{c}$ is the estimated threshold

Table 6. Out-of-sample forecast evaluation results

	Linear	Nonlinear (threshold variable)	
		(t)	$(\Delta_4 y_{t-9})$
MSPE $(k=1)$	5.11	4.48	**3.67**
MAPE $(k=1)$	1.83	1.56	**1.48**
S $(k=1)$		-1.58	-1.58
DM $(k=1)$		-0.61	-2.07*
MSPE $(k=2)$	3.47	**1.82**	2.04
MAPE $(k=2)$	1.48	**1.03**	1.06
S $(k=2)$		-2.21*	-3.44*
DM $(k=2)$		-2.39*	-2.05*
MSPE $(k=3)$	2.88	**1.52**	1.73
MAPE $(k=3)$	1.40	**0.91**	1.00
S $(k=3)$		-3.10*	-2.60*
DM $(k=3)$		-2.35*	-2.01*
MSPE $(k=4)$	2.23	**1.33**	1.38
MAPE $(k=4)$	1.20	**0.88**	0.93
S $(k=4)$		0.00	-0.75
DM $(k=4)$		-1.85	-1.67
MSPE $(k=5)$	1.93	1.16	**1.03**
MAPE $(k=6)$	1.11	0.83	**0.81**
S $(k=5)$		-0.62	-2.39*
DM $(k=5)$		-1.44	-2.26*
MSPE $(k=6)$	1.73	1.12	**0.96**
MAPE $(k=6)$	1.07	0.81	**0.78**
S $(k=6)$		-1.52	-3.55*
DM $(k=6)$		-1.30	-2.59*

Notes: The minimum value within each row for the MSPE and MAPE criteria is in bold type. * Indicates that the forecasts reject the null hypothesis of equality at the 5% significance level.

Table 7. Probit estimation results 1970:1 - 2000:4

$$\Pr[X_t = 1|z'_{t-k}, X_{t-k}] = F(\beta_0 + \beta_1 s_{t-3}^{EMU} + \beta_2 s_{t-12}^{US} + \gamma X_{t-4})$$

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}$	PR^2
Model (9)	-1.03	-1.27	-0.45		0.46
	(0.23)	(0.32)	(0.18)		
Model (10)	-0.99	-1.97	-0.59	-14.77	0.62
	(0.32)	(0.45)	(0.25)	(0.70)	

Notes: Numbers in parentheses report robust (QML) standard error
estimates. Robustness refers to the underlying distribution employed.
PR^2 is McFadden's pseudo R^2