CENTRE FOR APPLIED MACROECONOMIC ANALYSIS

The Australian National University

CAMA Working Paper Series **April, 2006**

NONLINEAR AUTOREGRESSIVE LEADING INDICATOR MODELS OF OUTPUT IN G-7 COUNTRIES

Heather M. Anderson
The Australian National University

George Athanasopoulos
Monash University

Farshid Vahid
The Australian National University

CAMA Working Paper 14/2006
http://cama.anu.edu.au

Nonlinear Autoregressive Leading Indicator Models of Output in G-7 Countries

Heather M. Anderson*
School of Economics and
Centre for Applied Macroeconomic Analysis
Australian National University
Canberra, ACT 0200
Australia

George Athanasopoulos
Department of Econometrics and Business Statistics
Monash University
Clayton, Victoria 3800
Australia

Farshid Vahid
School of Economics and
Centre for Applied Macroeconomic Analysis
Australian National University
Canberra, ACT 0200
Australia

*Corresponding author. E-mail: Heather.Anderson@anu.edu.au. Telephone: +61 2 6125 5439. Fax: +61 2 6125 5124

SUMMARY

This paper studies linear and nonlinear autoregressive leading indicator models of business cycles in G7 countries. Our models use the spread between short-term and long-term interest rates as leading indicators for GDP. We examine data admissibility by determining whether these models have the ability to produce time series with classical cycles that resemble the observed classical cycles in the data, and then we ask if this data admissibility lends itself to better predictions of the probability of recession.
Keywords: Business Cycles, Leading Indicators, Nonlinear Models, Data Admissibility, Probability Predictions, Model Evaluation.

JEL classification: C22, C23, E17, E37.

1. INTRODUCTION

Modeling the cyclical behavior of the aggregate output has always been an important question for macroeconomists, who often want to classify past and present patterns into particular phases of the business cycle, and forecast future turning points. There are lively debates about how to define and measure cycles in output, how to model them, and how to predict features such as turning points and recessions. Detrending issues fuel many of these debates (see e.g. Canova, 1998), but other important issues include possible nonlinearity in business cycles (see e.g. Hamilton, 1989 and Potter, 1995), and which variables are most useful for predicting output (Stock and Watson, 1989, and 2003).

The forecasting literature has often emphasized the ability of financial variables to predict various features of business cycles. In particular, Zellner and Hong (1989), Zellner *et al.* (1991) and Zellner and Min (1999) show that adding (lags of) monetary and financial variables to univariate autoregressive models of output growth improves forecasts of turning points in many countries. These authors call their models "autoregressive leading indicator" (ARLI) models, a term that we use from now on. Related to forecasts and ARLI models is a large set of macroeconomic papers that document and explain why specific financial variables have leading information for the business cycle[1]. In their comprehensive review of this work, Stock and Watson (2003) conclude

[1]A representative sample includes Davis and Fagan (1997), Estrella and Mishkin (1998), Friedman and Kuttner (1998), Gertler and Lown (1999), Hamilton and Kim (2002) and Kwark (2002).

that "there is evidence that the term spread is a serious candidate as a predictor of output growth and recessions. The stability of this proposition in the U.S. is questionable, however, and its universality is unresolved". We interpret the lack of stability in the output growth/term spread relationship as possible evidence of nonlinearity, and this motivates our nonlinear approach to modeling output and the spread.

Almost all bivariate analyses of output and the spread are based on linear specifications. However, the empirical finance literature presents statistically significant evidence that interest rate spreads are nonlinear (see e.g. Anderson, 1997, and Balke and Fomby, 1997). This suggests that a satisfactory bivariate model of output and the spread is likely to be nonlinear. Further, although the apparent decline in the variance of output growth in the United States since the mid eighties (see e.g. Kim and Nelson, 1999, and McConnell and Perez-Quiros, 2000) has led to a belief that it is necessary to include an exogenous structural change in the variance of output growth, it is possible that a nonlinear propagation mechanism might adequately generate changes in cycle characteristics, without any need for a structural break in variance.

In this paper we develop linear and nonlinear autoregressive leading indicator models of output growth in G-7 countries. Our models use the spread between short-term and long-term interest rates as leading indicators for growth in GDP, and we initially evaluate them according to their data admissibility, as gauged by the non-parametric procedures developed by Harding and Pagan (1999 and 2002). We then ask whether data admissible models can predict the probability of a recession (as in Fair, 1993) better than data inadmissible models. This contrasts with Teräsvirta and Anderson (1992), Clements and Krolzig (1998) and Jansen and Oh (1999), who used mean squared errors of one step ahead forecasts to evaluate various univariate nonlinear models of output. Our primary aim is to develop time series models that can predict important prespecified events such as "recessions" and other salient features of business cycles. We are interested in comparing the predictive ability of nonlinear specifications relative to linear specifications, and where applicable, relative to univariate linear models that incorporate a structural break. We are also interested in whether our data admissibility criteria have any bearing on a model's ability to predict.

Data admissibility is a property that is often required from a good model. Hendry (1995, p 364) states "A model is data admissible if its predictions automatically satisfy all known data constraints", but then he points out that although inadmissible models are logically inconsistent with the data, this does not necessarily mean that they are poor empirical representations. Data admissible models need not outperform all inadmissible models with respect all goodness of fit measures, and data inadmissible models may sometimes produce better point forecasts (for example, despite being data inadmissible due to zero returns, models of log squared returns in Harvey, Ruiz and Shephard 1994, and Engle and Marcucci 2006 produce better volatility forecasts than models of squared returns). Here, we want to study this issue in a business cycle context.

We use the procedures in Harding and Pagan (1999, 2002) to assess whether a model is "admissible with respect to business cycle characteristics" (calling this "data admissible" from now on). These procedures ask if a model's generated time series reflect characteristics of historical classical cycles, such as average durations and amplitudes of recessions and expansions, and measures of cycle shape. Harding and Pagan (1999) find that univariate linear and non-linear models for US, UK and Australian output growth rates have difficulty satisfying this requirement, as do linear vector autoregressions (VARs) of growth rates of output, consumption and investment. Here, we consider if multivariate autoregressive leading indicator models are data admissible, and if so, whether this means that they can predict probabilities of recessions better than non-admissible models.

In line with results of Stock and Watson (1989), Davis and Fagan (1997), Kozicki (1997), Friedman and Kuttner (1998), Estrella and Mishkin (1998) and others, we use yield spreads as our leading indicators. The predictive power of the spread is well established, but most research on this issue has stayed within the confines of conventional linear models of output and the spread. Notable exceptions include Estrella and Mishkin (1998) and Birchenhall *et al.* (2000) who use a logit/probit model to explain a binary recession indicator. Galbraith and Tkacz (2000) test for and find asymmetries in the link between the yield spread and output in G7 countries, but here we go a step further by explicitly modeling these asymmetries. Our work also extends work done by De Long and Summers (1988), Cover (1992), Karras (1996), Choi (1999) and Weise (1999) who model asymmetries between monetary policy and output, because we use the spread (rather than monetary growth) as an indicator of monetary policy.

We consider two classes of nonlinear models that each incorporate regimes that can easily be interpreted as recessionary and expansionary states. One of these classes contains nonlinear ARLI (NARLI) models, such as those studied in Anderson and Vahid (2001). These models are threshold and smooth transition autoregressive specifications, that allow an observed business cycle indicator to generate changes in regime. The other class of nonlinear models is the set of Markov switching VAR (MSVAR) models, introduced by Krolzig (1997), that allow an unobserved exogenous state variable to generate regime changes. Clements and Krolzig (2004) ask whether MSVAR models are admissible models for business cycle analysis, but only do so within a system of real macroeconomic aggregates.

We find that bivariate nonlinear models of output and the interest rate spread are often data admissible, whereas many univariate linear models and models with structural breaks are not. Thus, nonlinearity in a leading indicator framework appears to be important for building data admissible models. Further, our forecasting statistics provide some support for the proposition that data admissible models produce better predictions of probability of recession. For the USA, Canada and the UK, where a decline in volatility is observed, fixed parameter nonlinear leading indicator models perform better than univariate models with an exogenous structural break. For the other countries, the nonlinearity in the bivariate framework offers a noticeable improvement compared to univariate models, but a small improvement relative to bivariate VARs.

The next section of this paper describes our modeling methodology and develops the linear and nonlinear models that we use in our analysis. In Section 3, we discuss how we assess data admissibility and how we form our predictions of the probability of recession, and then we evaluate our models according to these criteria. We conclude in Section 4, with a summary of our findings and some directions for future research.

2. MODELING METHODOLOGY

2.1. Data

Our data consists of quarterly time series of real output (gross domestic product), short term interest rates and long term interest rates for the United States, Canada, the United Kingdom, France, Germany, Italy and Japan, and we base our benchmark analyses of business cycle characteristics on the natural logarithms of real GDP. Our spread variables are calculated by taking the difference between the interest rates on the long-term bond and the short term bond, and the variables in our parametric models are output growth (calculated as 100 $\times$ the differenced logarithms of real GDP) and the spread. Changes in the range of bonds offered in the last four countries have prevented the compilation of continuous spread series beyond the turn of the century, but continuous series are available for the United States, Canada and the United Kingdom, and our analysis on these three countries is therefore more extensive. We use the notation y_t to denote output growth (which we will call output) and s_t to denote the interest rate spread. We provide detailed information on data sources, our samples, and precise descriptions of our raw series in Appendix 1.

2.2. Linear and Nonlinear model specification

We develop our models, one country at a time, to make sure that we account for country specific characteristics. In each case, we estimate a univariate autoregressive specification, and then a VAR in output and the spread to provide a baseline bivariate model. We use AIC[2] to guide our lag-length choices, but eliminate lagged variables if they are statistically insignificant and their removal does not lead to serially correlated residuals. We estimate our restricted VARs both equation by equation (with OLS) and as a SUR, but there is never much difference between the two and we work with the latter. The output equation in the restricted VAR is an ARLI model. We also estimate random walk models for each country, so that later we can compare the simulated properties of the data and random walk models with other linear models, and thereby assess how the lag structure and the financial indicator in each ARLI model can account for data admissibility. We also test for structural breaks using Hansen's (1997) test, and build corresponding break models when they are supported by these tests.

We start the development of our nonlinear models by conducting nonlinearity tests on (i) an unrestricted linear model of output and (ii) each unrestricted equation in the VAR model. Our initial alternatives are threshold autoregressive (TAR) and logistic smooth transition autoregressive (LSTAR) models, because the relevant tests are easy to conduct and can be expected to also have power against Markov switching (MS) alternatives. A univariate LSTAR(p) model with transition variable y_{t-d} is defined by

$$y_t = (\pi_{10}+\pi_1' w_t)+(\pi_{20}+\pi_2' w_t)F(y_{t-d})+u_t, \quad \text{with } F(y_{t-d}) = [1+\exp[-\gamma(y_{t-d}-c)]\,]^{-1},$$

for $w_t = (y_{t-1}, \ldots, y_{t-p})'$, $y_{t-d} \in w_t$, $\pi_j = (\pi_{j1}, \ldots, \pi_{jp})'$ for $j = 1, 2$ and $u_t \sim nid$, and a TAR(p) model with transition variable y_{t-d} is simply the limiting case when $\gamma \to \infty$. Bivariate generalizations of these models are straightforward.

We use the Tsay (1989) test against the TAR alternative, and three tests by Luukkonen *et al.* (1988) and Teräsvirta (1994) for evidence of STAR behavior. All of these

[2] AIC tends to overparameterize and thereby minimize residual serial correlation that might interfere with our subsequent linearity tests.

tests require specifying the transition variable in advance. The Tsay (1989) test orders the data matrix (dependent and lagged variables) according to the transition variable, recursively estimates the model, and then tests whether the recursive residuals are orthogonal to the regressors. The three tests proposed by Luukkonen *et al.* (1988) are all tests for omitted nonlinear terms. The simplest of these, which we call the "first order test", takes the cross product of the transition variable and all regressors to be the omitted variables. The "augmented first order test" adds the third power of the transition variable to the list of cross products considered by the first order test, and the "third order test" uses the cross products of the first, second and third powers of the transition variable with all the regressors. Luukkonen *et al.* (1988) discuss the relative merits of these tests. As suggested by these authors, we use F-test versions of the tests to account for the relatively small sample size.

$<$ Insert Table 1 about here $>$

We consider each lag of output as a possible transition variable for univariate models of output, and consider each lag of output and the spread in bivariate cases. We perform nonlinearity tests for each country and summarize the results in Table 1. For each country, this table shows if there is evidence of nonlinearity in the univariate autoregressive model of output, in the output equation of a bivariate model of output and the spread, and in the spread equation of the bivariate model. An entry like y_{t-2} means that the null of linearity was rejected at the 5% level of significance when the transition variable was y_{t-2}. After the transition variable, we report which test or tests rejected linearity. The letters 'F', 'A', 'S' and 'T' stand for first order, augmented first order, third order and TAR tests respectively. Finally, the transition variable that we have selected for our final specification for each equation is marked by a star superscript.

The null of linearity is not rejected in the univariate autoregressive models of output for the US, Canada, UK and France. This is consistent with work in Anderson and Vahid (2001), who found no evidence of nonlinearity in the growth of US output (and then showed that fitting a univariate nonlinear autoregressive models to this data did not improve model performance relative to a linear AR model). However, in the bivariate setting, we find significant evidence of nonlinearity in both output and spread equations for all countries, except in the output equation for France.

We require statistically significant evidence of (S)TAR nonlinearity to proceed with any subsequent estimation of a corresponding (S)TAR model, and if evidence of nonlinearity is found for more than one transition variable, we fit separate nonlinear models for each transition variable, and then choose between (reduced versions) of estimated specifications by considering fit and whether models pass tests of no residual serial correlation. Since TAR models are special cases of STAR models when $\gamma \to \infty$, we begin by fitting a STAR model in all cases where nonlinearity is found, but we monitor the likelihood function, and switch to a TAR specification if the global maximum seems to occur when the transition parameter is very large.

We use the evidence of nonlinearity found in the bivariate tests for each country to justify the estimation of MSVAR models. A two state MSVAR(q) model for the vector of output growth and the spread given by $z_t \equiv (y_t, s_t)'$, is[3]

$$z_t = c(\kappa_t) + \Phi(\kappa_t) z_{t-1} + \cdots + \Phi(\kappa_t) z_{t-q} + u_t, \qquad u_t | \kappa_t \sim NID(0, \Sigma(\kappa_t)),$$

where κ_t is the latent state of the system that has two possible values, and the evolution of the state is governed by the matrix of transition probabilities

$$P = \begin{pmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{pmatrix}.$$

Given the large number of parameters of general MSVAR models relative to the size of our samples, we use the Schwarz criterion to choose the lag length, the number of regimes, and whether or not to allow the error variance covariance matrix to depend on the state. The estimations are performed using the MSVAR package (Krolzig, 1997) with the console version of Ox (Doornik, 2002).

For each country we estimate a random walk with drift model, a univariate linear autoregressive model (if different from the random walk), a univariate non-linear autoregressive model (if warranted by nonlinearity tests), a univariate AR with a structural break (if warranted by a structural break test), a bivariate linear model of output and the spread, a bivariate Markov switching model and a bivariate threshold or smooth transition model of output and the spread. Non-linear models are fitted only if tests of linearity reject the linear autoregressive models and if these models allow for sufficient number of observations in each regime. Since the univariate models are easily reproducible, we only report the estimated bivariate models for each country. These can be found in Appendix 2. Residual tests indicate that none of the dynamic models

[3] Another version of MSVAR models specifies the variables in mean subtracted form. We do not consider this form here, because it complicates the simulation of recession probabilities.

have serially correlated residuals (according to Lagrange multiplier tests), but many of the linear specifications show evidence of nonlinearity as indicated in Table 1, evidence of heteroskedasticity (according to White tests and ARCH tests), and evidence of structural change (according to Ramsey reset tests).

3. MODEL EVALUATION

We evaluate our models along two different dimensions, focussing first on whether they are data admissible, and then on their ability to predict recessions. This places a direct emphasis on the needs of model users, who will typically want to study and forecast business cycles. Of course, true business cycles are unobservable, but each evaluation technique is based on a well accepted way of classifying our output series into different phases of the business cycle. Both methods of classifications used in our evaluation are "model-free", and are popular in the business cycle literature.

3.1. Determining data admissibility with respect to business cycle characteristics

Harding and Pagan (2002) point out the gap between policy makers' focus on turning points in the levels of output and academic interest in modeling the moments of detrended data. They advocate using a cycle dating algorithm to identify the turning points in the levels, and measuring various business cycle characteristics (BCCs) based on these turning points. These BCCs include the duration and amplitude of a cycle from peak to trough and from trough to peak, as well as cumulative movements and asymmetries within these phases. We follow their suggested techniques for dating cycles and measuring eight BCCs, and then we evaluate our models by comparing the BCCs in our samples with the range of BCCs predicted by our models. This model evaluation technique can be seen as a test for model admissibility, in the sense that it tells us whether a model is likely to generate a feature that is actually observed in the data. A brief summary of this procedure is provided below.

The cycle dating algorithm is an adaptation of the Bry-Boschan (1971) algorithm, and it identifies turning points when

$$\log GDP_t > (<) \log GDP_{t\pm k} \quad \text{for } k = 1, 2 \text{ quarters},$$

provided that each phase of a cycle lasts at least two quarters and the whole cycle lasts at least five quarters. This algorithm is applied to both raw data, and data series that have been generated using DGPs implied by our estimated models.

Figure 1 illustrates the measurement of four BCCs over a peak to trough phase, while the economy moves along the curved path from point X to point Y. The length of the line XZ shows the duration of the phase, i.e. how long it takes (in quarters) for the phase to be completed, while the length of the line YZ shows the amplitude of the phase, i.e. the total change in output as the economy moves from X to Y. We convert the latter into a percentage change. The shaded area labeled "cumulation" measures the impact of the recession, by approximating the total accumulated loss in output as the economy moves from peak to trough. We convert this measure to a percentage. The final BCC (labeled "excess" in the right hand diagram in Figure 1) measures the difference between the cumulated output loss and a crude triangle approximation (given by triangle XYZ) to this loss. This measures the curvature of the phase of the cycle. We divide this measure by the duration and convert it to a percentage.

< Insert Figure 1 about here >

The above measures relate to a single recession, but one can summarize the business cycle characteristics of a given series by calculating the means of each BCC for all peak-to-trough and all trough-to-peak phases. These eight summary statistics (calculated without any prior detrending of the series) provide a natural benchmark for evaluating a business cycle model, because a data admissible model should have the potential to reproduce the BCCs that are actually observed in the data. We simulate growth series from our estimated models and then integrate the simulated series to obtain an analogue of the original data together with its BCC measures. For each parametric model, we undertake 10000 simulations in order to estimate the model dependent density functions for each of the eight characteristics of interest, and then we compare these densities with the relevant characteristics in the original data. If an observed BCC lies in the upper 5% or lower 5% tails of the simulated density, then we conclude that the parametric model is unlikely to produce data with the observed business cycle characteristics, and is therefore an inadmissible model. It is important to note that we use Harding and Pagan's procedure as a data admissibility criterion, rather than as a measure of goodness of fit or a test. We ask if a model is likely to ever produce a time series with the observed BCCs, and not how closely it fits the observed BCCs.

Nonlinear models will generally produce time series that have a wider range of BCCs than linear models, simply because these models are less restrictive. Therefore, in cases where linear models are data admissible, nonlinear models are expected to be data admissible as well. We are interested in those cases where linear models are found to be too restrictive. We want to know if the bivariate nonlinear models that we develop here are general enough to satisfy the data admissibility criteria, and whether this is important for predicting the probability of recession.

3.2. Predicting probabilities of recessions

In line with previous work done by Neftçi (1982), Diebold and Rudebusch (1989), Zellner *et al.*. (1991) and Fair (1993), we evaluate models according to their ability in predicting business cycle events. Prediction is based on simulation, since some of our models are nonlinear and taking expectations is not straightforward. The technique for predicting the probability of an event involves defining the event of interest as a property of a sequence of multi-step ahead predictions, classifying each predicted sequence from the simulations as either having or not having that property, and then setting the estimated probability equal to the proportion of simulated sequences that have the property. See Fair (1993) for further details.

Fair (1993) uses two definitions of a recession, which are:

A: At least two *consecutive* quarters of negative growth in real GDP over the next five quarters; and

B: At least two quarters of negative growth in real GDP over the next five quarters.

Although we have evaluated all models according to their ability to predict each of these events, we only report the results for event A because it corresponds to a definition of a recession that is in common use. Overall conclusions based on predictions of event B are qualitatively similar[4]. We take lagged observations and our estimated parameters as given for each observation in our sample, and then use the simulation process to estimate the probabilities of events A and B. This leads to series of probabilities (P_t), which can be compared against indicator variables (D_t) for events A and B, where each of P_t and D_t relate to an event over the five quarter period between (t) and ($t+4$), and P_t is predicted from the information set at time ($t-1$).

[4]The results for event B are available from the authors upon request.

It is useful to note the similarities and differences between our probability predictions, and others. Firstly, we define a recession as an observable event. Therefore, we do not need to make inference about an unobservable state. Also, since our definition of recession is directly related to one to five period ahead forecasts of output, an appropriate model for forecasting the probability of recession is one that delivers the one to five period ahead predictive density of output. In this context, binary dependent variable models are problem specific, and if there is interest in estimating the probabilities associated with other events, then the dependent variable needs to be redefined in each case, and the model needs to be re-estimated.

We evaluate our probability forecasts using Brier's (1950) quadratic probability scores (QPS) and log probability scores (LPS), which are respectively defined by

$$QPS = \frac{1}{T}\sum_{t=1}^{T} 2(P_t - D_t)^2 \qquad (0 < QPS < 2), \quad \text{and}$$

$$LPS = -\frac{1}{T}\sum_{t=1}^{T} [(1 - D_t)\ln(1 - P_t) + D_t \ln P_t] \qquad (0 < LPS < \infty)$$

for a sample of T forecasts. QPS provides a probability analogue to the usual mean squared error criterion, while LPS penalizes large mistakes more than QPS. Like the mean squared error measure, low QPS and low LPS imply accurate forecasts. See Diebold and Rudebusch (1989) for further discussion on these evaluation criteria. Since in all cases that a model had noticeably lower QPS than other models it had noticeably lower LPS as well, we report only the QPS results.[5]

Out-of-sample evaluation is often problematic when forecasting rare events, and in our context, the fact that there have been no consecutive quarters of negative growth in the US, Canada and the UK during the last 10 years, makes a short out-of-sample evaluation uninformative. We therefore track model performance over the *entire* sample, rather than over a short post-sample evaluation period. Since QPS and LPS criteria differ from the loss functions that are minimized when the parameters are estimated, there is little reason to believe that they necessarily improve with the fit of the model. Indeed, our results show that larger models do not necessarily outperform more parsimonious models. For the US, Canada and the UK, we also present the QPS scores for out-of-sample forecasts relating to 2000:1 to 2004:4, although we place little emphasis on these because of our concern that this period is not representative.

[5] The LPS scores for all models are available from the authors upon request.

3.3. Results

Summary statistics of the business cycle characteristics for each of our log(GDP) series are provided in Table 2. Here, it is clear that each characteristic varies from country to country, and that the characteristics of the peak-to-trough phase are quite different from those of the trough-to-peak phase. Of course, given that the log(GDP) series have a positive trend and that the table shows the characteristics of the actual series (as opposed to the detrended series), it is not surprising that the trough-to-peak duration and amplitude of the cycles are much larger than the peak-to-trough ones. One striking observation is the small trough-to-peak characteristics, in particular the cumulative gain, of Japanese business cycles relative to other countries. We attribute this to the clear break in the trend in Japanese GDP since 1990. As we will discuss below, we cannot reproduce this break endogenously with a nonlinear self-exciting model.

< Insert Table 2 about here >

Summary tables of the performance of different models in reproducing the business cycle characteristics for each of the G-7 countries are reported in Tables 3A to 3G, and quadratic probability scores of different models in predicting recessions for all seven countries are reported in Table 4. We highlight the main findings for each country below.

< Insert Tables 3A to 3G and Tables 4 and 5 about here >

The United States: Using data from 1947:1 to 1997:1, Harding and Pagan (2002) find that linear autoregressive models are able to reproduce all characteristics of cycles in the US except its curvature during expansions. Our analysis based on more recent data shows that a univariate AR(2) model can produce all business cycle characteristics, as can our bivariate linear and nonlinear leading indicator models. The quadratic probability scores in the first column of Table 4 show that there is a noticeable improvement in predictions of recession probabilities when using leading indicators, although there is not much difference between linear and non-linear bivariate models. Regarding the relationship between data admissibility and ability to predict recessions, the US case provides little information. The random walk with drift model cannot produce cycles with peak to trough amplitude and cumulation similar to the actual data, and it produces the worst probability forecasts. A univariate model that allows for break in the variance and the autoregressive structure in 1984:2 has difficulty in reproducing trough to peak curvature, and it produces worse probability forecasts than the bivariate models for the period 1961:1 to 1999:4. Not surprisingly, the break model performs the best for the "out of sample" period 2000:1 to 2004:4 (see Table 5) simply because this period did not contain two consecutive quarters of negative growth.

Canada: The Canadian results are more informative about the relationship between data admissibility and the ability to predict recessions. Those models that failed the evaluations of data admissibility in more dimensions, performed more poorly when predicting the probability of recession. The only models that passed our data admissibility criteria in all dimensions were the two bivariate nonlinear specifications, and these two models produced the most accurate predictions of the probability of recession. As in the US case, most of the ability to predict results from adding the leading indicator.

United Kingdom: The results for the United Kingdom are similar to those for Canada in that the only models that passed our data admissibility criteria in all dimensions were the two bivariate nonlinear specifications, and these two models produced the most accurate predictions of the probability of recession. The differences in abilities to predict are not as pronounced as those for Canada, and most of the ability to predict results from adding nonlinearity to the leading indicator model.

France: The French case is the only case where no nonlinearity is found in either the univariate output equation or in the ARLI model of output. Moreover, in the MSVAR class of models, although the Schwarz criterion chooses a two-state homoskedastic MSVAR(1), this criterion prefers the linear VAR(1) model. Only the random walk model fails to meet the data admissibility criteria, and this model is one of the worst with respect to predicting the probability of recession. All bivariate models produce better probability predictions than univariate models.

Germany: This is the first case where the univariate autoregressive model of output shows significant signs of nonlinearity, and hence we have also estimated a univariate LSTAR model for output. However, the univariate nonlinear model does not produce better probability forecasts than the univariate linear autoregressive model. All models (except the random walk model) satisfy data admissibility criteria. Bivariate models score better in predicting the probability of recessions, with bivariate nonlinear models improving the scores only slightly over the bivariate linear model.

Italy: As in the German case, linearity is rejected even in the linear autoregressive model of output, and all estimated models satisfy data admissibility criteria. However, unlike the German case, the univariate nonlinear model of output scores considerably better than the univariate linear model in predicting recessions. The contribution of the spread to the output equation is quite weak and the QPS show that the addition of the spread does not help predict recessions in Italy. The univariate nonlinear model performs almost as well as the NARLI model.

Japan: The Japanese case is unique in the sense that without allowing for an exogenous structural break, all models fail to capture the shape of the business cycles in Japan. We attribute this to the fact that in Japan there has been a significant decline in the output trend, unlike the US or the UK cases where the evidence of a break is in the variance and the persistence of output growth. We estimate the break date to be 1991:2. Linear models with the break can capture the shape of the business cycles in Japan. The ARLI model with break passes all data admissibility criteria and produces the best forecasts for the probability of recessions. We emphasize that in the forecasting exercise, we assume that the structural break is recognized immediately after it happens.

4. CONCLUSION AND DIRECTIONS FOR FURTHER RESEARCH

In this paper we ask if bivariate nonlinear autoregressive models of output growth and the term spread can explain and forecast important features of business cycles in G-7 countries. We evaluate our models by assessing whether or not they are admissible models of business cycle, and how well they can forecast the probability of well defined events such as "two consecutive quarters with negative growth in the next five quarters". We use the measures proposed by Harding and Pagan (2002) as our criteria for data admissibility, and we use a simulation method used by Fair (1993) to produce each model's prediction of recession probabilities.

Our results shows some evidence of correlation between data admissibility and ability to forecast recessions. The majority of models satisfy the admissibility criteria. There are cases like Canada and the UK where the bivariate nonlinear models are the only models that pass the data admissibility criteria and they also produce the best probability predictions. Also, in no country are the best probability predictions produced by a data inadmissible model.

A clearer picture emerges about the usefulness of considering bivariate models of output growth and the spread for predicting recessions. In all countries other than Italy, the bivariate models of output and spread produce better predictions of recession probabilities than univariate models. We consider two types of bivariate nonlinear models that allow for transition between two distinct dynamic types of behavior that can be interpreted as expansionary and recessionary dynamics. Both types of nonlinear models do well in predicting recessions, with the addition of nonlinearity to the UK bivariate leading indicator models leading to a particularly large improvement. Future research might consider working with larger systems that include more leading indicators.

Relative to other research that has looked at the link between yield spreads and output, the distinctive feature of our work is that we explicitly model nonlinearities in output and the spread. We follow a stepwise procedure for developing our models which starts from finding the best linear model and then moves to nonlinear models only if nonlinear models are warranted by the data.

We have only allowed for exogenous structural breaks in our univariate models, with the exceptional case of Japan where, without a break in the mean growth, no bivariate model could pass the data admissibility criteria. This allows us to conclude that for countries other than Japan, exogenous structural breaks are not essential for developing models that are data admissible and can produce good predictions of recessions. An avenue for future research would be to establish if allowing for exogenous structural breaks in bivariate models of output and the spread further improves the ability of these models in predicting recession probabilities in countries such as the US, the UK and Canada. This opens up interesting possibilities such as different breaks in each equation or same breaks in both equations (co-breaking). At the same time, because samples are small and recessions are rare events, a belief in frequent unexplained exogenous structural breaks could cast doubt on the usefulness of time series models for forecasting recessions.

Some researchers have advocated the use of error correction terms from models of financial markets in models of real variables, and we think that this insight is important. The yield spread is, of course a valid error correction term in modeling the bond market, and as such, it summarizes many features of the bond sector. Recent related work by Sensier *et al.* (2002) provides evidence of the usefulness of short-term interest rates in Germany for predicting recession in Italy and France. We believe that further research that uses carefully chosen error correction terms from international financial markets as predictors for output within a multivariate nonlinear framework, may lead to superior models for capturing the shape and the turning points of business cycles.

ACKNOWLEDGEMENTS:
We acknowledge financial support from the Australian Research Council (Large Grant No. A00103258), and research assistance from Jing Tian and Pawin Siriprapanukul.

References

Anderson HM. 1997. Transaction costs and nonlinear adjustment towards equilibrium in the US treasury bill market. *Oxford Bulletin of Economics and Statistics* **59**: 465-484.

Anderson HM, Vahid F. 2001. Predicting the probability of a recession with nonlinear autoregressive leading indicator models. *Macroeconomic Dynamics* **5**: 482-505.

Balke NS, Fomby TB. 1997. Threshold cointegration. *International Economic Review* **38**: 627-647.

Birchenhall CR, Jessen H, Osborn DR, Simpson P. 1999. Predicting U.S. business cycle regimes. *Journal of Business and Economic Statistics* **17**: 313-323.

Brier GW. 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* **75**: 1-3.

Bry G, Boschan C. 1971. *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. New York. NBER.

Canova F. 1998. Detrending and business cycle facts. *Journal of Monetary Economics* **41**: 475-512.

Choi WG. 1999. Asymmetric monetary effects on interest rates across monetary policy stances. *Journal of Money, Credit and Banking* **31**: 386-416.

Cover JP. 1992. Asymmetric effects of positive and negative money-supply shocks. *Quarterly Journal of Economics* **107**: 1261-1282.

Clements MP, Krolzig H. 2004. Can regime switching models reproduce the business cycle features of the united states?. *International Journal of Finance and Economics* **9**: 1-14.

Clements MP, and Krolzig H. 1998. A comparison of the forecast performance of Markov switching and threshold autoregressive models of US GNP. *Econometrics Journal* **1**: c47-c75.

Davis EP, Fagan G. 1997. Are financial spreads useful indicators of future inflation and output growth in EU countries?. *Journal of Applied Econometrics* **12**: 701-714.

De Long BJ, and Summers L. (1988). How does macroeconomic policy affect output?. *Brookings Papers on Economic Activity* **2**: 433-480.

Diebold FX, Rudebusch GD. 1989. Scoring the leading indicators. *Journal of Business* **62**: 369-391.

Doornik JA. 2002. *Object Oriented Matrix Programming Using OX.* http://www.doornik.com.

Engle RF, Marcucci J. 2006. A long-run pure variance common feature model for the common volatilities of the Dow Jones. Forthcoming in the *Journal of Econometrics.*

Estrella A, Mishkin FS. 1998. Predicting U.S. recessions: financial variables as leading indicators. *Review of Economics and Statistics* **80**: 45-61.

Fair RC. 1993. Estimating event probabilities from macroeconometric models using stochastic simulation. In *Business Cycles, Indicators and Forecasting*, Stock J, Watson M. (eds). NBER: Chicago.

Friedman BM, Kuttner KN. 1998. Indicator properties of the paper-bill spread: lessons from recent experience. *Review of Economics and Statistics* **80**: 34-44.

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.

Gertler M, and Lown CS. 1999. The information in the high-yield bond spread for the business cycle: evidence and some implications. *Oxford Review of Economic Policy* **15**: 132-150.

Hamilton JD. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**: 357-381.

Hamilton JD, Kim DH. 2002. A Re-examination of the predictability of economic activity using the yield spread. *Journal of Money, Credit and Banking* **34**: 340-360.

Hansen BE. 1997. Approximate asymptotic p-values for structural change tests. *Journal of Business and Economic Statistics* **15**: 60-67.

Harding A, Pagan AR. 1999. Dissecting the cycle. Melbourne Institute Working Paper No. 13/99.

Harding A, Pagan AR. 2002. Dissecting the cycle: a methodological investigation. *Journal of Monetary Economics* **49**: 365-381.

Harvey AC, Ruiz E, Shephard N. 1994. Multivariate stochastic variance models. *Review of Economic Studies* **61**: 247-264.

Hendry DF. 1995. *Dynamic Econometrics.* Oxford University Press: Oxford.

Jansen DW, and Oh W. 1999. Modeling nonlinearity of business cycles: choosing between the CDR and STAR models. *Review of Economics and Statistics* **81**: 344-349.

Karras G. 1996. Are the output effects of monetary policy asymmetric? Evidence from a sample of European countries. *Oxford Bulletin of Economics and Statistics* **58**: 267-278.

Kim C, Nelson CR. 1999. Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics* **81**: 608-616.

Kozicki S. 1997. Predicting real growth and inflation with the yield spread. *Federal Reserve Bank of Kansas City Economic Review* 39-57.

Krolzig HM. 1997. *Markov-Switching Vector Autoregressive Modelling: Statistical Inference and Applications to Business Cycle Analysis.* Springer Verlag.

Kwark N-S. 2002. Default risk, interest rate spreads, and business cycles: explaining the interest rate spread as a leading indicator. *Journal of Economic Dynamics and Control* **26**: 271-302.

Luukkonen R, Saikkonen P, Teräsvirta T. 1988. Testing linearity against smooth transition autoregressive models. *Biometrika* **75**: 491-499.

McConnell MM, and Perez-Quiros G. 2000. Output fluctuations in the United States: what has changed since the early 1980s?. *American Economic Review* **90**: 1464-1476.

MSVAR 1.3 for Ox. http://www.economics.ox.ac.uk/research/hendry/krolzig/index.html.

Neftçi SN. 1982. Optimal prediction of cyclical downturns. *Journal of Economic Dynamics and Control* **4**: 225-241.

Potter SM. 1995. A nonlinear approach to US GDP. *Journal of Applied Econometrics* **10**: 109-125.

Sensier M, Artis M, Osborn DR, Birchenhall CR. 2004. Domestic and international influences on business cycle regimes in Europe. *International Journal of Forecasting* **20**: 343-357.

Stock JH, Watson MW. 1989. New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*. NBER.

Stock JH, Watson MW. 2003. Forecasting output and inflation: the role of asset prices. *Journal of Economic Literature* 788 - 829.

Teräsvirta T. Anderson HM. 1992. Characterizing nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics* **7**: S119-S136.

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

Tsay R. 1989. Testing and modeling threshold autoregressive processes. *Journal of the American Statistical Association*. **84**: 231-240.

Weise CL. 1999. The asymmetric effects of monetary policy: a nonlinear vector autoregressive approach. *Journal of Money, Credit and Banking* **31**: 85-108.

Zellner A, Hong C. 1989. Forecasting international growth rates using Bayesian shrinkage and other procedures. *Journal of Econometrics* **40**: 183-202.

Zellner A, Hong C, Min C. 1991. Forecasting turning points in international growth rates using Bayesian exponentially weighted autoregression, time-varying parameter and pooling techniques. *Journal of Econometrics* **49**: 275-304.

Zellner A, Min C. 1999. Forecasting turning points in countries' output growth rates. *Journal of Econometrics* **88**: 203-306.

APPENDIX 1: DATA

Precise descriptions of the raw series that we use in this analysis are given below. Unless otherwise stated, we have extracted all data for the last three countries from the data files available on Mark Watson's web page. We use the logarithms of real GDP when we undertake our benchmark analysis, and our models are functions of output growth ($y_t = 100 \times \Delta \ln(GDP)$) and the interest rate spread ($s_t =$ Long-term interest rate - Short-term interest rate). The effective samples used for analysis are shorter than the raw series because of lagged variables in the models.

USA (1960:1 to 2004:4)

Output: Real Gross Domestic Product: (Billions of Chained 2000 Dollars, seasonally adjusted at annual rates, from the U.S. Federal Reserve Economic Data Base (FRED)).

Short-Term Interest Rates: 3-Month Treasury (Secondary) Bill Market Rates (Averages over business days expressed as a percentage, from FRED).

Long-Term Interest Rates: 10-Year Treasury Bond Constant Maturity Rates (Averages over business days expressed as a percentage, from FRED).

The effective sample consisted of 154 observations, dating from 1961:1 to 1999:4. The out of sample forecast period ran from 2000:1 - 2004:4.

CANADA (1961:1 to 2004:3)

Output: Real Gross Domestic Product (seasonally adjusted in constant 1997 prices, from the Bank of Canada).

Short-Term Interest Rates: Interest rates on 90 day deposit receipts. (expressed as a percentage pa, series CAN.IRT3DR01.ST from the DX database (Australia)).

Long-Term Interest Rates: Yields on long term government bonds (>10 Years). (expressed as a percentage pa, series CAN.IRLGV06.ST from the DX database (Australia)).

The effective sample consisted of 153 observations, dating from 1961:4 to 1999:4. The out of sample forecast period ran from 2000:1 - 2004:3.

UNITED KINGDOM (1960:1 to 2004:4)

Output: Real Gross Domestic Product (seasonally adjusted chained volume index with 2001=100, from the Office of National Statistics).

Short-Term Interest Rates: 3 Month Treasury Bill Rates. (expressed as a percentage pa, series 11260C..ZF... from the IFS portion of the DX database (Australia)).

Long-Term Interest Rates: Yields on 10 Year Government Bonds (expressed as a percentage pa, series GBR.IRLTGV02.ST from the DX database (Australia)).

The effective sample consisted of 158 observations, dating from 1960:3 to 1999:4. The out of sample forecast period ran from 2000:1 - 2004:4.

FRANCE (1970:1 to 1998:4)

Output: Real Gross Domestic Product (seasonally adjusted in constant 1980 prices, series FRA.NAGVTT01h.NCALSA).

Short-Term Interest Rates: Interest Rate on 3 Month PIBOR (expressed as a percentage pa, series FRA.IRT31B01.ST).

Long-Term Interest Rates: Interest Rates on 10 year Bonds when issued (expressed as a percentage pa, series FRA.IRLTOT02.ST).

The effective sample consisted of 113 observations, dating from 1970:4 to 1998:4.

GERMANY (1960:1 to 1999:4)

Output: Real Gross Domestic Product, (seasonally adjusted, series I 199bv&r@c134, originally from the IFS data base).

Short-Term Interest Rates: Overnight Interest Rate (expressed as a percentage pa, series I 160c@c134, from the IFS data base).

Long-Term Interest Rates: Interest rate on a long term Government bond (expressed as a percentage pa, series I 161@c134, from the IFS data base).

The effective sample consisted of 155 observations, dating from 1961:2 to 1999:4.

ITALY (1971:4 to 1998:4)

Output: Real Gross Domestic Product, (seasonally adjusted, series I 199bv&r@c136, originally from the IFS data base).

Short-Term Interest Rates: Overnight Interest Rate (expressed as a percentage pa, series I 160c@c136, from the IFS data base).

Long-Term Interest Rates: Interest Rate on a Long Term Government Bond (expressed as a percentage pa, series I 161@c136, from the IFS data base).

The effective sample consisted of 110 observations, dating from 1972:3 to 1998:4.

JAPAN (1969:4 to 1999:4)

Output: Real Gross Domestic Product, (seasonally adjusted in 1990 prices, series I 199bv&r@c158, originally from the IFS data base).

Short-Term Interest Rates: Overnight Interest Rate (expressed as a percentage pa, series I 160b@c158, from the IFS data base).

Long-Term Interest Rates: Interest Rate on a Long Term Government Bond (expressed as a percentage pa, series I 161@c158, from the IFS data base).

The effective sample consisted of 113 observations, dating from 1971:2 to 1999:4.

APPENDIX 2: BIVARIATE MODELS OF OUTPUT AND SPREAD

A. USA: 1961:3 - 1999:4

(Standard errors are in brackets)

ARLI model of output and spread:

$$\hat{y}_t = \underset{(0.12)}{0.32} + \underset{(0.08)}{0.18} y_{t-1} + \underset{(0.08)}{0.14} y_{t-2} + \underset{(0.06)}{0.20} s_{t-2} \qquad \hat{\sigma}_{\text{MLE}} = 0.79$$

$$\hat{s}_t = \underset{(0.09)}{0.45} - \underset{(0.05)}{0.13} y_{t-2} - \underset{(0.05)}{0.19} y_{t-5} + \underset{(0.08)}{1.03} s_{t-1} - \underset{(0.11)}{0.35} s_{t-2} + \underset{(0.08)}{0.19} s_{t-3} \qquad \hat{\sigma}_{\text{MLE}} = 0.53$$

NARLI model of output and spread:

$$\begin{aligned} \hat{y}_t &= \underset{(0.22)}{-0.66} + \underset{(0.07)}{0.22} y_{t-1} - \underset{(0.16)}{0.26} y_{t-3} + \underset{(0.18)}{0.56} y_{t-4} + \underset{(0.13)}{0.30} y_{t-5} + \underset{(0.15)}{0.94} s_{t-2} - \underset{(0.12)}{0.30} s_{t-3} \\ & \quad f_{yt} \times (\underset{(0.27)}{1.56} + \underset{(0.18)}{0.33} y_{t-3} - \underset{(0.18)}{0.56} y_{t-4} - \underset{(0.15)}{0.53} y_{t-5} - \underset{(0.12)}{0.63} s_{t-2}) \\ f_{yt} &= 1 * (y_{t-2} > 0.32) \qquad \hat{\sigma}_{\text{MLE}} = 0.69 \end{aligned}$$

$$\begin{aligned} \hat{s}_t &= \underset{(0.09)}{0.73} - \underset{(0.06)}{0.20} y_{t-1} - \underset{(0.05)}{0.13} y_{t-2} - \underset{(0.07)}{0.20} y_{t-4} - \underset{(0.06)}{0.28} y_{t-5} + \underset{(0.10)}{0.82} s_{t-1} - \underset{(0.14)}{0.59} s_{t-2} + \underset{(0.11)}{0.67} s_{t-3} \\ & \quad f_{yt} \times (\underset{(0.07)}{0.20} y_{t-4} + \underset{(0.08)}{0.17} y_{t-5} + \underset{(0.14)}{0.40} s_{t-1} + \underset{(0.20)}{0.34} s_{t-2} - \underset{(0.14)}{0.82} s_{t-3}) \\ f_{yt} &= 1 * (y_{t-1} > 0.73) \qquad \hat{\sigma}_{\text{MLE}} = 0.45 \end{aligned}$$

MSVAR model of output and spread:

$$\text{Regime 1} \quad : \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.17)}{0.09} \\ \underset{(0.20)}{0.38} \end{pmatrix} + \begin{pmatrix} \underset{(0.15)}{0.02} & \underset{(0.15)}{-0.12} \\ \underset{(0.17)}{0.10} & \underset{(0.17)}{0.91} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix} +$$

$$\begin{pmatrix} \underset{(0.13)}{0.08} & \underset{(0.16)}{0.58} \\ \underset{(0.15)}{-0.06} & \underset{(0.19)}{-0.13} \end{pmatrix} \begin{pmatrix} y_{t-2} \\ s_{t-2} \end{pmatrix}, \qquad \Sigma_1 = \begin{pmatrix} 0.65 & -0.23 \\ -0.23 & 0.90 \end{pmatrix}$$

$$\text{Regime 2} \quad : \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.17)}{0.67} \\ \underset{(0.07)}{0.32} \end{pmatrix} + \begin{pmatrix} \underset{(0.09)}{0.13} & \underset{(0.20)}{0.32} \\ \underset{(0.04)}{-0.17} & \underset{(0.08)}{1.34} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix} +$$

$$\begin{pmatrix} \underset{(0.10)}{0.17} & \underset{(0.20)}{-0.30} \\ \underset{(0.04)}{-0.07} & \underset{(0.09)}{-0.43} \end{pmatrix} \begin{pmatrix} y_{t-2} \\ s_{t-2} \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0.51 & 0.01 \\ 0.01 & 0.08 \end{pmatrix}$$

$$\hat{P}_{11} = 0.88, \ \hat{P}_{22} = 0.96.$$

B. Canada: 1961:1 - 1999:4

(Standard errors are in brackets)

ARLI model of output and spread:

$$\hat{y}_t = \underset{(0.10)}{0.50} + \underset{(0.08)}{0.17} y_{t-1} + \underset{(0.04)}{0.21} s_{t-2} \qquad \hat{\sigma}_{\text{MLE}} = 0.80$$

$$\hat{s}_t = \underset{(0.09)}{0.30} - \underset{(0.07)}{0.21} y_{t-1} + \underset{(0.04)}{0.90} s_{t-1} \qquad \hat{\sigma}_{\text{MLE}} = 0.74$$

NARLI model of output and spread:

$$\begin{aligned} \hat{y}_t &= \underset{(0.19)}{0.74} y_{t-2} + \underset{(0.10)}{0.59} s_{t-1} + \\ &\quad f_{yt} \times (\underset{(0.13)}{0.54} + \underset{(0.09)}{0.23} y_{t-1} - \underset{(0.19)}{0.74} y_{t-2} - \underset{(0.10)}{0.59} s_{t-1} + \underset{(0.05)}{0.12} s_{t-2}) \\ f_{yt} &= (1 + \exp\{-21.21 y_{t-1}\})^{-1} \qquad \hat{\sigma}_{\text{MLE}} = 0.73 \end{aligned}$$

$$\begin{aligned} \hat{s}_t &= \underset{(0.16)}{1.18} + \underset{(0.09)}{0.66} s_{t-2} \\ &\quad f_{st} \times (-\underset{(0.16)}{1.18} + \underset{(0.09)}{1.19} s_{t-1} - \underset{(0.12)}{0.94} s_{t-2}) \\ f_{st} &= (1 + \exp\{-138.08\,(y_{t-1} - 0.02)\})^{-1} \qquad \hat{\sigma}_{\text{MLE}} = 0.66 \end{aligned}$$

MSVAR model of output and spread:

$$\text{Regime 1:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.09)}{0.10} \\ \underset{(0.15)}{0.34} \end{pmatrix} + \begin{pmatrix} \underset{(0.11)}{0.56} & \underset{(0.04)}{0.10} \\ \underset{(0.17)}{-0.53} & \underset{(0.06)}{0.91} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 0.32 & -0.06 \\ -0.06 & 0.32 \end{pmatrix}$$

$$\text{Regime 2:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.24)}{1.12} \\ \underset{(0.14)}{0.33} \end{pmatrix} + \begin{pmatrix} \underset{(0.11)}{-0.07} & \underset{(0.10)}{0.13} \\ \underset{(0.06)}{-0.11} & \underset{(0.07)}{0.86} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0.71 & 0.02 \\ 0.02 & 0.23 \end{pmatrix}$$

$$\hat{P}_{11} = 0.93, \ \hat{P}_{22} = 0.95.$$

C. UK: 1961:1 - 1999:4

(Standard errors are in brackets)

ARLI model of output and spread:

$$\hat{y}_t = \underset{(0.10)}{0.49} + \underset{(0.05)}{0.12}\, s_{t-3} \qquad \hat{\sigma}_{\text{MLE}} = 1.03$$

$$\hat{s}_t = \underset{(0.07)}{0.20} + \underset{(0.07)}{1.15}\, s_{t-1} - \underset{(0.08)}{0.26}\, s_{t-2} - \underset{(0.05)}{0.14}\, y_{t-3} \qquad \hat{\sigma}_{\text{MLE}} = 0.68$$

NARLI model of output and spread:

$$\begin{aligned} \hat{y}_t &= \underset{(0.11)}{0.35} + \underset{(0.09)}{0.21}\, y_{t-2} + \underset{(0.08)}{0.15}\, y_{t-3} + \underset{(0.11)}{0.32}\, s_{t-1} + f_{yt} \times (-\underset{(0.24)}{0.51}\, s_{t-1}) \\ f_{yt} &= (1 + \exp\{-2.67\,(y_{t-2} - 0.68)\})^{-1} \qquad \hat{\sigma}_{\text{MLE}} = 0.98 \end{aligned}$$

$$\hat{s}_t = \underset{(0.07)}{0.20} - \underset{(0.05)}{0.14}\, y_{t-3} + \underset{(0.08)}{1.14}\, s_{t-1} - \underset{(0.08)}{0.26}\, s_{t-2} \qquad \hat{\sigma}_{\text{MLE}} = 0.68$$

MSVAR model of output and spread:

$$\text{Regime 1} \;:\; \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.14)}{0.40} \\ \underset{(0.11)}{-0.13} \end{pmatrix} + \begin{pmatrix} \underset{(0.10)}{0.00} & \underset{(0.10)}{0.19} \\ \underset{(0.08)}{0.00} & \underset{(0.09)}{1.39} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix} +$$

$$\begin{pmatrix} \underset{(0.08)}{0.30} & \underset{(0.10)}{-0.12} \\ \underset{(0.08)}{0.16} & \underset{(0.09)}{-0.48} \end{pmatrix} \begin{pmatrix} y_{t-2} \\ s_{t-2} \end{pmatrix}, \qquad \Sigma_1 = \begin{pmatrix} 0.18 & -0.04 \\ -0.04 & 0.14 \end{pmatrix}$$

$$\text{Regime 2} \;:\; \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.20)}{0.45} \\ \underset{(0.15)}{0.29} \end{pmatrix} + \begin{pmatrix} \underset{(0.12)}{-0.08} & \underset{(0.20)}{-0.02} \\ \underset{(0.08)}{0.00} & \underset{(0.13)}{1.02} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix} +$$

$$\begin{pmatrix} \underset{(0.14)}{-0.11} & \underset{(0.20)}{0.18} \\ \underset{(0.10)}{-0.10} & \underset{(0.13)}{-0.18} \end{pmatrix} \begin{pmatrix} y_{t-2} \\ s_{t-2} \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 1.71 & -0.10 \\ -0.10 & 0.70 \end{pmatrix}$$

$$\hat{P}_{11} = 0.79,\; \hat{P}_{22} = 0.81.$$

D. France: 1970:4 - 1998:4

(Standard errors are in brackets)

ARLI model of output and spread:

$$\hat{y}_t = \underset{(0.08)}{0.35} + \underset{(0.09)}{0.19} y_{t-2} + \underset{(0.04)}{0.13} s_{t-2} \qquad \hat{\sigma}_{\text{MLE}} = 0.58$$

$$\hat{s}_t = \underset{(0.10)}{0.22} + \underset{(0.06)}{0.80} s_{t-1} \qquad \hat{\sigma}_{\text{MLE}} = 0.91$$

NARLI model of output and spread:

$$\hat{y}_t = \underset{(0.08)}{0.34} + \underset{(0.09)}{0.19} y_{t-2} + \underset{(0.04)}{0.14} s_{t-2} \qquad \hat{\sigma}_{\text{MLE}} = 0.58$$

$$\begin{aligned} \hat{s}_t &= -\underset{(0.38)}{1.64} + \underset{(0.49)}{0.58} y_{t-2} + \\ (s_{t-1} &> -0.33) \times \left(\underset{(0.41)}{2.18} - \underset{(0.49)}{0.58} y_{t-2} + \underset{(0.10)}{1.01} s_{t-1} - \underset{(0.09)}{0.38} s_{t-2} \right) \qquad \hat{\sigma}_{\text{MLE}} = 0.78 \end{aligned}$$

MSVAR model of output and spread:

$$\text{Regime 1:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.20)}{0.27} \\ \underset{(0.25)}{-1.54} \end{pmatrix} + \begin{pmatrix} \underset{(0.20)}{0.57} & \underset{(0.09)}{0.19} \\ \underset{(0.23)}{0.94} & \underset{(0.11)}{0.63} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 0.32 & -0.02 \\ -0.02 & 0.32 \end{pmatrix}$$

$$\text{Regime 2:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.10)}{0.36} \\ \underset{(0.11)}{0.85} \end{pmatrix} + \begin{pmatrix} \underset{(0.10)}{0.08} & \underset{(0.05)}{0.12} \\ \underset{(0.11)}{-0.13} & \underset{(0.05)}{0.62} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_2 = \Sigma_1$$

$$\hat{P}_{11} = 0.72, \; \hat{P}_{22} = 0.91.$$

E. Germany: 1961:2 - 1999:4

(Standard errors are in brackets)

ARLI model of output and spread:

$$\hat{y}_t = \underset{(0.14)}{0.26} - \underset{(0.07)}{0.23} y_{t-1} + \underset{(0.07)}{0.30} y_{t-4} + \underset{(0.06)}{0.23} s_{t-3} \qquad \hat{\sigma}_{\text{MLE}} = 1.17$$

$$\hat{s}_t = \underset{(0.10)}{0.36} + \underset{(0.06)}{0.82} s_{t-1} + \underset{(0.09)}{0.24} s_{t-3} - \underset{(0.08)}{0.29} s_{t-4} \qquad \hat{\sigma}_{\text{MLE}} = 0.87$$

NARLI model of output and spread:

$$\begin{aligned} \hat{y}_t &= \underset{(0.42)}{0.99} y_{t-2} + \underset{(0.35)}{1.80} y_{t-3} + \underset{(0.06)}{0.18} s_{t-1} + (1 + \exp\{-0.90\,(y_{t-1} + 1.4)\})^{-1} \times \\ & \left(\underset{(0.16)}{0.50} - \underset{(0.54)}{1.21} y_{t-2} - \underset{(0.42)}{2.24} y_{t-3}\right) \qquad \hat{\sigma}_{\text{MLE}} = 1.06 \end{aligned}$$

$$\begin{aligned} \hat{s}_t &= \underset{(0.07)}{0.81} s_{t-1} + \underset{(0.68)}{1.43} s_{t-2} - \underset{(0.58)}{1.17} s_{t-4} + (1 + \exp\{-1.41\,(y_{t-3} + 0.90)\})^{-1} \times \\ & \left(\underset{(0.12)}{0.33} - \underset{(0.72)}{1.59} s_{t-2} + \underset{(0.10)}{0.21} s_{t-3} + \underset{(0.62)}{1.06} s_{t-4}\right) \qquad \hat{\sigma}_{\text{MLE}} = 0.78 \end{aligned}$$

MSVAR model of output and spread:

$$\text{Regime 1:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.18)}{0.12} \\ \underset{(0.43)}{0.33} \end{pmatrix} + \begin{pmatrix} \underset{(0.14)}{0.48} & \underset{(0.08)}{0.14} \\ \underset{(0.31)}{-0.11} & \underset{(0.18)}{0.64} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 0.57 & -0.28 \\ -0.28 & 3.18 \end{pmatrix}$$

$$\text{Regime 2:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.17)}{0.55} \\ \underset{(0.08)}{0.22} \end{pmatrix} + \begin{pmatrix} \underset{(0.10)}{-0.37} & \underset{(0.08)}{0.24} \\ \underset{(0.04)}{-0.07} & \underset{(0.03)}{0.90} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 1.56 & 0.05 \\ 0.05 & 0.28 \end{pmatrix}$$

$$\hat{P}_{11} = 0.83, \; \hat{P}_{22} = 0.97.$$

F. Italy: 1971:3 - 1999:4

(Standard errors are in brackets)

ARLI model of output and spread:

$$\hat{y}_t = \underset{(0.08)}{0.36} + \underset{(0.08)}{0.43} y_{t-1} + \underset{(0.03)}{0.09} s_{t-2}, \qquad \hat{\sigma}_{\mathrm{MLE}} = 0.73$$

$$\hat{s}_t = \underset{(0.16)}{0.48} + \underset{(0.09)}{0.87} s_{t-1} - \underset{(0.12)}{0.30} s_{t-2} + \underset{(0.09)}{0.17} s_{t-3} - \underset{(0.15)}{0.49} y_{t-1} - \underset{(0.15)}{0.40} y_{t-3}, \qquad \hat{\sigma}_{\mathrm{MLE}} = 1.25$$

NARLI model of output and spread:

$$\hat{y}_t = \underset{(0.08)}{0.43} + \underset{(0.10)}{0.35} y_{t-1} + (1 + \exp\{-6.56\,(y_{t-3} - 1.40)\})^{-1} \times \left(\underset{(0.29)}{0.81} y_{t-1} - \underset{(0.15)}{0.37} y_{t-3} + \underset{(0.09)}{0.30} s_{t-2}\right) \qquad \hat{\sigma}_{\mathrm{MLE}} = 0.65$$

$$\hat{s}_t = \underset{(0.17)}{0.39} + \underset{(0.09)}{0.78} s_{t-1} - \underset{(0.11)}{0.32} s_{t-2} + \underset{(0.08)}{0.18} s_{t-3} - \underset{(0.14)}{0.48} y_{t-1} - \underset{(0.16)}{0.57} y_{t-3} + (1 + \exp\{-7.47\,(s_{t-1} - 2.0)\})^{-1} \times \left(\underset{(2.13)}{7.63} - \underset{(0.80)}{2.85} s_{t-1} + \underset{(0.41)}{0.82} y_{t-3}\right) \quad \hat{\sigma}_{\mathrm{MLE}} = 1.13$$

MSVAR model of output and spread:

Regime 1:
$$\begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.27)}{0.44} \\ \underset{(0.60)}{-0.48} \end{pmatrix} + \begin{pmatrix} \underset{(0.18)}{0.77} & \underset{(0.08)}{0.10} \\ \underset{(0.37)}{-1.14} & \underset{(0.17)}{0.55} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 0.91 & 1.06 \\ 1.06 & 3.56 \end{pmatrix}$$

Regime 2:
$$\begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.09)}{0.40} \\ \underset{(0.12)}{0.25} \end{pmatrix} + \begin{pmatrix} \underset{(0.09)}{0.26} & \underset{(0.04)}{0.03} \\ \underset{(0.11)}{-0.16} & \underset{(0.05)}{0.90} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0.34 & -0.01 \\ -0.01 & 0.47 \end{pmatrix}$$

$$\hat{P}_{11} = 0.51, \ \hat{P}_{22} = 0.87.$$

G. Japan: 1971:2 - 1999:4

(Standard errors are in brackets)

ARLI model of output and spread (with break):

$$\hat{y}_t = \underset{(0.15)}{1.04} - \underset{(0.19)}{0.90} D_t + \underset{(0.09)}{0.15} y_{t-3} - \underset{(0.08)}{0.18} y_{t-5} + \underset{(0.06)}{0.15} s_{t-1} \qquad \hat{\sigma}_{\text{MLE}} = 0.82$$

$$\hat{s}_t = \underset{(0.08)}{0.14} + \underset{(0.06)}{0.89} s_{t-1} - \underset{(0.05)}{0.17} s_{t-4} \qquad \hat{\sigma}_{\text{MLE}} = 0.75$$

NARLI model of output and spread:

$$\begin{aligned} \hat{y}_t &= \underset{(0.11)}{0.47} + \underset{(0.08)}{0.29} y_{t-3} + \underset{(0.08)}{0.21} y_{t-4} + \underset{(0.11)}{0.22} s_{t-1} - \underset{(0.09)}{0.19} s_{t-2} + \\ &\quad f_{yt} \times \left(-\underset{(0.53)}{1.10} + \underset{(0.41)}{0.85} y_{t-2} - \underset{(0.28)}{0.93} s_{t-4} \right) \\ \widehat{f}_{yt} &= \left(1 + \exp\left\{-12.69\left(y_{t-5} - 2.00\right)\right\}\right)^{-1} \qquad \hat{\sigma}_{\text{MLE}} = 0.77 \\ \hat{s}_t &= \underset{(0.08)}{0.14} + \underset{(0.06)}{0.89} s_{t-1} - \underset{(0.05)}{0.17} s_{t-4} \qquad \hat{\sigma}_{\text{MLE}} = 0.75 \end{aligned}$$

MSVAR model of output and spread:

$$\text{Regime 1:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.18)}{1.20} \\ \underset{(0.21)}{-0.16} \end{pmatrix} + \begin{pmatrix} \underset{(0.14)}{-0.22} & \underset{(0.09)}{0.41} \\ \underset{(0.16)}{-0.06} & \underset{(0.10)}{0.91} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 0.59 & 0.17 \\ 0.17 & 0.84 \end{pmatrix}$$

$$\text{Regime 2:} \quad \begin{pmatrix} \hat{y}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \underset{(0.16)}{0.69} \\ \underset{(0.06)}{0.48} \end{pmatrix} + \begin{pmatrix} \underset{(0.15)}{0.04} & \underset{(0.11)}{-0.07} \\ \underset{(0.05)}{0.17} & \underset{(0.04)}{0.57} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ s_{t-1} \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0.84 & 0.05 \\ 0.05 & 0.12 \end{pmatrix}$$

$$\hat{P}_{11} = 0.93, \ \hat{P}_{22} = 0.94.$$

Table 1: Evidence of Nonlinearity

	USA	Canada	UK	France	Germany	Italy	Japan
Univariate	-	-	-	-	y^*_{t-1} FAST	y^*_{t-3} AST	y_{t-4} T
Models of					y_{t-2} FAST		y^*_{t-5} T
Output					y_{t-3} FAST		
					y_{t-4} S		
Bivariate	y_{t-1}S	y^*_{t-1} FAST	y^*_{t-2} FAST		y^*_{t-1} FAST	y^*_{t-3} AS	y^*_{t-4} S
Models of	y^*_{t-2} FAST	s_{t-1} FA	y_{t-3} FS		y_{t-2} FAS	s_{t-1} AT	s_{t-1} S
Output	y_{t-5} T		s_{t-1} S		y_{t-3} FAST	s_{t-2} FA	
	s_{t-1}FAST		s_{t-2} S				
	s_{t-2}FAT						
	s_{t-3}FAST						
Bivariate	y^*_{t-1} AST	y^*_{t-1} FAST		y_{t-1} FAST	y_{t-2} S	y_{t-3} S	y_{t-2} S
Models of	y_{t-2} FAST	s_{t-1} ST		y_{t-2} FAST	y^*_{t-3} FAST	s^*_{t-1} S	s_{t-1} S
Spread	y_{t-3} FAS			s^*_{t-1} FAST	s_{t-1} T		s^*_{t-4} FAS
	y_{t-4}S			s_{t-2} FAST	s_{t-2} FAST		s_{t-5} T
	y_{t-5} T				s_{t-3} FAS		
	s_{t-1} FAST				s_{t-4} FA		
	s_{t-2} ST						
	s_{t-3} FAST						

Entries in the table relate to rejections of the null hypothesis of linearity against threshold and smooth transition autoregressive alternatives at the 5% significance level. The stated lag of the growth rate (y) or the spread (s) is the transition variable in the alternative specification. The superscripts F, A, S and T relate to four alternative linearity tests. See the text for descriptions of the tests.

Table 2: Benchmark Business Cycle Characteristics

	USA	Canada	UK*	France	Germany	Italy	Japan
Time Span	61:1-99:4	61:4-99:4	60:1-99:4	70:4-98:4	61:2-99:4	71:3-99:4	71:2-99:4
Duration							
PT	3.2	4.0	3.8	3.0	4.5	2.8	3.6
TP	21.0	16.0	19.6	32.5	19.2	14.8	8.0
Amplitude							
PT	-2.0	-2.9	-2.6	-1.6	-2.3	-1.5	-1.9
TP	22.9	17.0	17.2	21.3	20.1	11.5	4.9
Cumulation							
PT	-3.1	-6.5	-6.9	-2.0	-5.3	-3.0	-6.3
TP	351	251	258	358	253	130	19
Excess							
PT	-0.03	0.19	-0.04	0.03	0.11	-0.06	-0.13
TP	1.24	1.21	0.07	-0.37	0.79	0.27	-0.13

Note: The UK figures relate to a cycle with a minimum length of 4 quarters rather than 5. This makes our analysis comparable to Harding and Pagan (1999).

Table 3A: Simulated Business Cycle Characteristics for the United States

	Raw Data	RW + Drift	AR(2)	AR(2)+ Break[a]	ARLI[b]	MSVAR[b]	NARLI[b]
Duration							
PT	3.2	2.5 (2.0,3.5)	3.1 (2.0,4.7)	3.7 (2.4,5.6)	3.1 (2.0,4.7)	2.9 (2.0,4.5)	4.0 (2.2,6.3)
TP	21.0	34.3 (13.4,72.0)	29.3 (13.0,60.0)	19.8 (10.5,35.0)	31.1 (16.4,60.5)	38.4 (11.5,92.0)	33.5 (16.3,67.5)
Amplitude							
PT	−2.0	−1.1* (−1.7,−0.5)	−1.4 (−2.4,−0.7)	−2.4 (−4.1,−1.2)	−1.4 (−2.4,−0.7)	−1.5 (−2.9,−0.5)	−2.5 (−5.1,−0.8)
TP	22.9	31.9 (12.3,67.2)	29.3 (12.0,61.9)	22.2 (10.9,41.3)	30.9 (16.3,58.7)	40.3 (11.5,96.0)	33.5 (17.8,68.0)
Cumulation							
PT	−3.1	−1.5* (−2.9,−0.5)	−2.8 (−6.5,−0.8)	−6.4 (−17.2,−1.6)	−2.7 (−6.3,−0.8)	−2.7 (−7.7,−0.5)	−6.9 (−18.6,−0.9)
TP	351	912 (105,2952)	762 (108,2395)	394 (84,1082)	785 (183,2278)	1258 (77,4466)	967 (193,2942)
Excess							
PT	−0.03	0.00 (−0.1,0.1)	0.00 (−0.2,0.2)	0.00 (−0.2,0.2)	0.00 (−0.2,0.2)	0.00 (−0.2,0.2)	0.08 (−0.2,0.4)
TP	1.24	0.04 (−1.4,1.5)	0.03 (−1.4,1.5)	−0.05* (−1.2,1.0)	0.11 (−1.1,1.4)	0.07 (−2.0,2.2)	0.36 (−0.8,1.6)

[a] : The break is estimated to be at 1983:3.

[b] : Full specifications of all bivariate models are given in Appendix B.

* : The observed characteristic is outside of the 90% simulated band.

Table 3B: Simulated Business Cycle Characteristics for Canada

	Raw Data	RW + Drift	AR(1)	AR(1)+ Break[a]	ARLI[b]	MSVAR[b]	NARLI[b]
Duration							
PT	4.0	2.4* (2.0,3.5)	2.8 (2.0,4.0)	3.8 (2.4,5.8)	3.3 (2.0,5.0)	3.6 (2.0,5.6)	3.5 (2.0,5.5)
TP	16.0	35.0 (0.0,84.0)	28.2 (12.3,57.5)	15.9 (7.8,30.0)	24.9 (10.4,46.0)	28.4 (9.3,63.0)	33.6 (10.5,80.0)
Amplitude							
PT	−2.9	−1.1* (−1.8,−0.5)	−1.3* (−2.2,−0.7)	−2.5 (−4.3,−1.3)	−1.6* (−2.7,−0.8)	−1.7 (−3.0,−0.7)	−2.4 (−5.3,−0.7)
TP	17.0	34.9 (0.0,84.1)	29.3 (12.3,60.3)	18.0 (7.8,35.3)	23.9 (9.6,44.2)	29.9 (7.8,72.6)	34.7 (9.8,84.9)
Cumulation							
PT	−6.5	−1.4* (−2.9,−0.4)	−2.2* (−4.7,−0.7)	−6.8 (−16.9,−1.8)	−3.3 (−7.8,−0.9)	−4.0 (−10.3,−0.8)	−5.9 (−17.8,−0.7)
TP	251	1062 (0,3718)	717 (101,2220)	261 (42,760)	465 (68,1233)	811 (45,2821)	990 (61,3593)
Excess							
PT	0.19	0.00* (−0.1,0.1)	0.00* (−0.1,0.1)	0.00 (−0.2,0.2)	0.00 (−0.2,0.2)	0.01 (−0.2,0.2)	0.06 (−0.2,0.3)
TP	1.21	0.05 (−1.5,1.7)	0.04 (−1.4,1.5)	0.07* (−1.0,1.1)	0.25 (−1.2,1.9)	0.09 (−1.6,1.8)	0.10 (−1.6,1.9)

[a] : The break is estimated to be at 1974:2.

[b] : Full specifications of all bivariate models are given in Appendix B.

* : The observed characteristic is outside of the 90% simulated band.

Table 3C: Simulated Business Cycle Characteristics for UK

	Raw Data	RW + Drift	AR(1) + Break[a]	ARLI[b]	MSVAR[b]	NARLI[b]
Duration						
PT	3.8	3.0 (2.3,4.0)	3.6 (2.6,4.8)	3.1 (2.3,4.2)	3.3 (2.3,4.8)	3.5 (2.3,5.0)
TP	19.6	15.2 (8.8,25.6)	10.8* (6.8,16.7)	15.0 (8.8,25.0)	18.1 (9.4,32.7)	17.4 (9.4,30.8)
Amplitude						
PT	-2.6	−1.7* (−2.3,−1.1)	−2.8 (−4.0,−1.9)	−1.8* (−2.4,−1.2)	−2.2 (−3.2,−1.4)	−1.9 (−3.1,−1.1)
TP	17.2	12.9 (7.6,21.3)	11.7 (7.3,18.1)	12.9 (7.5,21.1)	13.8 (7.2,24.9)	15.1 (8.0,27.0)
Cumulation						
PT	-6.9	−2.8* (−5.1,−1.4)	−6.4 (−12.5,−2.9)	−3.2* (−5.9,−1.5)	−4.4 (−8.8,−1.7)	−4.6 (−11.0,−1.4)
TP	258	173 (48,427)	107 (36,249)	171 (48,413)	228 (49,621)	239 (56,648)
Excess						
PT	-0.04	0.00 (−0.1,0.1)	0.00 (−0.2,0.2)	0.00 (−0.1,0.1)	−0.03 (−0.3,0.2)	0.00 (−0.2,0.2)
TP	0.07	0.00 (−0.6,0.6)	−0.02 (−0.6,0.5)	0.01 (−0.6,0.6)	0.05 (−0.7,0.9)	0.02 (−0.7,0.7)

[a] : The break is estimated to be at 1981:1.

[b] : Full specifications of all bivariate models are given in Appendix B.

* : The observed characteristic is outside of the 90% simulated band.

Table 3D. Simulated Business Cycle Characteristics for France

	Raw Data	RW + Drift	AR(2)	ARLI[a]	MSVAR[a]	NARLI[a]
Duration						
PT	3.0	2.5 (2.0,3.5)	3.2 (2.0,5.0)	3.1 (2.0,5.0)	4.1 (2.0,8.0)	3.8 (2.0,7.0)
TP	32.5	28.7 (8.7,68.0)	25.2 (9.0,56.0)	28.0 (8.0,66.0)	24.2 (8.3,55.0)	22.6 (6.0,54.0)
Amplitude						
PT	−1.6	−0.8* (−1.3,−0.4)	−1.1 (−1.9,−0.5)	−1.0 (−1.8,−0.4)	−2.6 (−8.0,−0.5)	−1.2 (−2.4,−0.5)
TP	21.3	19.2 (5.8,45.1)	17.5 (5.1,40.9)	19.0 (4.6,46.6)	16.9 (5.1,39.5)	14.6 (2.8,38.3)
Cumulation						
PT	−2.0	−1.1 (−2.2,−0.3)	−2.1 (−5.4,−0.5)	−1.9 (−4.9,−0.4)	−11.7 (−47.4,−0.5)	−3.4 (−10.5,−0.5)
TP	358	434 (27,1545)	367 (28,1257)	438 (22,1572)	334 (25,1160)	314 (10,1188)
Excess						
PT	0.03	0.00 (−0.1,0.1)	0.00 (−0.1,0.1)	0.00 (−0.1,0.1)	0.09 (−0.2,0.5)	0.00 (−0.1,0.1)
TP	−0.37	0.03 (−1.0,1.2)	0.02 (−1.2,1.2)	0.01 (−1.3,1.3)	−0.02 (−1.1,1.1)	0.03 (−1.1,1.1)

[a] : Full specifications of all bivariate models are given in Appendix B.

* : The observed characteristic is outside of the 90% simulated band.

Table 3E. Simulated Business Cycle Characteristics for Germany

	Raw Data	RW + Drift	AR(4)	STAR(4)[a]	ARLI[b]	MSVAR[b]	NARLI[b]
Duration							
PT	4.5	3.3* (2.4,4.4)	3.4 (2.3,5.1)	3.2 (2.0,5.0)	3.8 (2.3,5.8)	3.4 (2.3,4.9)	3.4 (2.2,5.2)
TP	19.2	13.9 (8.5,22.8)	18.5 (10.1,32.7)	18.2 (9.9,31.7)	19.1 (10.1,35.0)	17.8 (9.7,31.7)	20.2 (10.7,37.7)
Amplitude							
PT	-2.3	-2.4 $(-3.3,-1.6)$	-2.1 $(-3.1,-1.3)$	-2.1 $(-3.4,-1.1)$	-2.3 $(-3.5,-1.3)$	-2.1 $(-3.1,-1.3)$	-1.9 $(-3.3,-1.1)$
TP	20.1	14.0 (8.6,22.4)	16.5 (8.6,30.0)	16.5 (8.6,29.2)	17.3 (8.6,32.5)	16.1 (8.5,28.8)	17.4 (9.3,31.7)
Cumulation							
PT	-5.3	-4.6 $(-8.5,-2.1)$	-4.5 $(-9.8,-1.6)$	-4.3 $(-11.1,-1.2)$	-5.7 $(-12.9,-1.7)$	-4.3 $(-9.2,-1.6)$	-4.5 $(-11.8,-1.2)$
TP	253	161 (49,401)	270 (60,731)	264 (61,686)	294 (61,831)	255 (58,680)	311 (71,858)
Excess							
PT	0.11	0.00 $(-0.2,0.2)$	0.00 $(-0.2,0.2)$	0.05 $(-0.1,0.3)$	0.00 $(-0.2,0.2)$	-0.01 $(-0.2,0.2)$	0.00 $(-0.2,0.2)$
TP	0.79	0.01* $(-0.6,0.6)$	0.02 $(-0.8,0.9)$	0.35 $(-0.5,1.3)$	0.04 $(-0.9,1.0)$	0.02 $(-0.7,0.8)$	0.22 $(-0.6,1.0)$

[a] : The transition variable of the logistic smooth transition model is the first lag of the growth rate.

[b] : Full specifications of all bivariate models are given in Appendix B.

* : The observed characteristic is outside of the 90% simulated band.

Table 3F. Simulated Business Cycle Characteristics for Italy

	Raw Data	RW + Drift	AR(5)	STAR(5)[a]	ARLI[b]	MSVAR[b]	NARLI[b]
Duration							
PT	2.8	2.8 (2.0,4.0)	3.5 (2.4,4.7)	3.0 (2.3,3.9)	3.6 (2.4,5.0)	3.3 (2.0,5.0)	3.2 (2.3,4.5)
TP	14.8	20.6 (9.3,43.0)	15.6 (8.8,27.7)	16.0 (9.0,28.8)	15.5 (8.7,27.0)	19.0 (9.0,38.0)	15.2 (8.4,26.7)
Amplitude							
PT	-1.5	−1.2 (−1.9,−0.7)	−1.7 (−2.6,−0.9)	−1.5 (−2.4,−0.8)	−1.8 (−2.8,−1.0)	−1.6 (−3.0,−0.7)	−2.1 (−3.7,−1.0)
TP	11.5	15.2 (6.9,31.2)	13.3 (7.5,22.7)	13.2 (7.4,22.7)	13.3 (7.3,22.8)	14.9 (6.5,29.7)	12.2 (6.6,20.6)
Cumulation							
PT	-3.0	−2.0 (−4.1,−0.7)	−3.5 (−6.9,−1.3)	−2.3 (−4.1,−1.0)	−3.9 (−8.1,−1.3)	−3.5 (−9.0,−0.8)	−4.3 (−10.4,−1.3)
TP	130	264 (41,837)	158 (38,424)	157 (40,425)	159 (40,420)	235 (38,702)	135 (34,355)
Excess							
PT	-0.06	0.00 (−0.1,0.2)	0.00 (−0.1,0.2)	0.00 (−0.1,0.1)	0.00 (−0.2,0.2)	0.00 (−0.2,0.2)	−0.04 (−0.3,0.1)
TP	0.27	0.02 (−0.8,0.9)	0.01 (−0.7,0.7)	−0.18 (−1.0,0.6)	0.02 (−0.7,0.8)	−0.05 (−1.1,1.0)	−0.37 (−1.3,0.3)

[a] : The transition variable of the logistic smooth transition model is the third lag of the growth rate.

[b] : Full specifications of all bivariate models are given in Appendix B.

Table 3G. Simulated Business Cycle Characteristics for Japan

	Raw Data	RW + Drift+Break[a]	AR(5)+ Break[a]	TAR(5)[b]	ARLI +Break[a]	MSVAR[c]	NARLI[c]
Duration							
PT	3.6	3.5 (2.0,6.0)	3.6 (2.0,6.0)	3.6 (2.0,6.7)	4.2 (2.3,7.0)	3.6 (2.0,6.7)	4.1 (2.0,6.5)
TP	8.0	17.5 (4.0,42.0)	18.6 (4.5,43.5)	27.1 (8.0,66.0)	17.8 (5.0,38.5)	20.2* (8.7,42.7)	22.5* (9.7,47.0)
Amplitude							
PT	−1.9	−1.7 (−2.0,−0.6)	−1.7 (−3.0,−0.6)	−1.6 (−3.2,−0.6)	−2.2 (−3.0,−0.6)	−1.9 (−4.0,−0.8)	−5.3 (−12.2,−1.2)
TP	4.9	15.4* (2.4,40.5)	16.5 (2.6,42.7)	24.5* (5.2,63.1)	17.4 (3.1,39.8)	18.2 (6.8,39.6)	25.8* (8.7,55.8)
Cumulation							
PT	−6.3	−3.8 (−10.4,−0.8)	−4.0 (−10.6,−0.9)	−4.2 (−12.6,−0.6)	−6.4 (−16.4,−1.3)	−6.0 (−20.3,−0.9)	−17.2 (−51.4,−1.6)
TP	19	353 (5,1348)	383 (6,1473)	554* (23,2119)	355 (9,1244)	327* (40,1051)	440* (49,1448)
Excess							
PT	−0.13	0.00 (−0.2,0.2)	0.00 (−0.2,0.2)	0.00 (−0.2,0.2)	0.00 (−0.3,0.3)	0.01 (−0.2,0.2)	−0.25 (−1.2,0.4)
TP	−0.13	0.34 (−0.6,1.8)	0.36 (−0.7,1.9)	0.04 (−1.8,1.9)	0.28 (−0.8,1.7)	0.06 (−1.1,1.3)	−0.39 (−2.4,1.9)

[a] : The break is a break in the constant term only at 1991:2.

[b] : The transition variable of the threshold autoregressive model is the fourth lag of the growth rate.

[c] : Full specifications of all bivariate models are given in Appendix B.

* : The observed characteristic is outside of the 90% simulated band.

Table 4: Quadratic Probability Scores for Predictions of Event A

Model	USA	Canada	UK	France	Germany	Italy	Japan
Constant	0.236	0.208	0.278	0.136	0.318	0.380	0.250
RW	0.238	0.209	0.288	0.138	0.333	0.384	0.200
AR	0.227	0.201	0.288	0.151	0.327	0.340	0.199
AR+Break	0.224	0.183	0.271	-	-	-	0.196
(S)TAR	-	-	-	-	0.330	0.297	0.194
ARLI	0.165	0.103	0.271	0.113	0.285	0.367	***0.180**
MSVAR	**0.153**	0.099	**0.231**	**0.104**	**0.271**	0.366	0.218
NARLI	0.172	**0.090**	0.243	0.111	0.281	**0.289**	0.192
Prob(Event A)	0.136	0.118	0.167	0.073	0.199	0.243	0.147

Notes: A dash (-) indicates that the model was not supported by the relevant test and was therefore not estimated. The star (*) indicates that the model had an intercept break.

Table 5: Quadratic Probability Scores for Predictions of Event A (2000:1-2004:4)

Model	USA	Canada	UK
Constant	0.037	0.027	0.058
RW	0.021	0.015	0.113
AR	0.046	0.041	0.113
AR+Break	**0.007**	0.075	**0.028**
ARLI	0.030	0.011	0.159
MSVAR	0.010	0.011	0.061
NARLI	0.022	**0.005**	0.091

Figure 1: Calculation of Business Cycle Characteristics

(Duration, Amplitude, Cumulation and Excess)