Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models

ARTICLE *in* JOURNAL OF APPLIED ECONOMETRICS · DECEMBER 1992

Impact Factor: 1.76 · DOI: 10.1002/jae.3950070509 · Source: RePEc

Timo Terasvirta
Aarhus University

JOURNAL OF APPLIED ECONOMETRICS, VOL. 7, S119–S136 (1992)

CHARACTERIZING NONLINEARITIES IN BUSINESS CYCLES USING SMOOTH TRANSITION AUTOREGRESSIVE MODELS

T. TERÄSVIRTA

Research Institute of the Finnish Economy, Lönnrotinkatu 4 B, SF-00120 Helsinki, Finland and University of California, San Diego, La Jolla, CA 92093-0508, USA

AND

H. M. ANDERSON

University of California, San Diego, La Jolla, CA 92093-0508, USA

SUMMARY

During the past few years investigators have found evidence indicating that various time-series representing business cycles, such as production and unemployment, may be nonlinear. In this paper it is assumed that if the time-series is nonlinear, then it can be adequately described by a smooth transition autoregressive (STAR) model. The paper describes the application of these models to quarterly logarithmic production indices for 13 countries and 'Europe'. Tests reject linearity for most of these series, and estimated STAR models indicate that the nonlinearity is needed mainly to describe the responses of production to large negative shocks such as oil price shocks.

1. INTRODUCTION

The possible nonlinearity of business cycles is an old topic in economics. Mitchell (1927, pp. 330–334 and 407–412) discussed this topic and presented statistical evidence both in favour of and against the asymmetry of business cycles. Keynes (1936, p. 314) argued that the contractions in an economy are more violent but also more short-lived than the expansions, so that GNP follows an asymmetric cyclical process with upswings which are longer than the downturns. Nonlinear models are needed to describe the data-generating mechanisms of inherently asymmetric realizations, because linear autoregressive or ARMA models are only capable of generating realizations with symmetrical cyclical fluctuations.

The issue of nonlinearity of business cycles is important because it has clear implications on business cycle theory. If business cycles are inherently nonlinear, then theory leading to linear models of business cycles has to be given up as inadequate. On a more practical level, if the indicators for business cycles are nonlinear, then linear models used to describe them may yield forecasts that are inferior to those from nonlinear models.

Recent empirical literature on the nonlinearity of business cycles includes tests of the null hypothesis of linearity in various business cycle indicators against specified and nonspecified nonlinear alternatives. Many of these tests have failed to reject the null. There is also a literature which models the first differences of quarterly postwar seasonally adjusted US GNP using various (nonlinear) two-regime switching models, but this modelling is not usually

0883–7252/92/0SS119–18$14.00

Received May 1991
Revised August 1992

supported by linearity tests. An exception to this literature is a paper by Tiao and Tsay (1991), which rejects linearity against threshold autoregression, and then fits a two-regime threshold autoregressive (TAR) model (Tong, 1990) to the data.

This paper can be regarded as a continuation of work in Luukkonen and Teräsvirta (1991), who use Lagrange multiplier type tests to test linearity in quarterly international industrial output and unemployment series against well-specified alternatives. These authors are frequently able to reject the null of linearity. Here, linearity in business cycle indicators is tested against well-specified alternatives, and the implied nonlinear models are estimated and studied when the null is rejected. Questions of interest include whether business cycle indicators can be adequately described by a flexible class of univariate nonlinear time-series models, and what the fitted models can tell us about the dynamics of these series. The approach used here relies on the assumption that the error distribution is symmetric, so that the nonlinearity which is found by the tests is attributed to a parameterized model, rather than to a nonsymmetric error process.

The assumption that the economy can only be in two states is generalized by basing the analysis on the smooth transition autoregressive (STAR) model. This model allows the business cycle indicator to alternate between two distinct regimes which represent two different phases of the business cycle, but transition between these regimes can be smooth, so that there can be a continuum of states between extreme regimes. The two-regime TAR model is a special case of the STAR model, so that our analysis accounts for the possibility that there might be only two regimes.

We use industrial production indices from 13 OECD countries and Europe for this investigation. Fitting a nonlinear model to a single time-series may be interesting in itself, but is not necessarily convincing because of the possibility of overfitting and thereby drawing hasty conclusions. We reduce the risk of drawing spurious conclusions by using a well-defined family of nonlinear models and data from several countries to look for similarities across countries in the estimated models. Model diagnostics and forecasting statistics are also examined as a means of evaluating the estimated nonlinear specifications.

The plan of the paper is as follows. We introduce our main tool, the STAR model, in Section 2. Section 3 is devoted to the specification of STAR models. The international industrial production data and the specification, estimation and evaluation of STAR models are discussed in Section 4. Section 5 compares forecasts from linear AR and STAR models and Section 6 concludes.

2. THE BASIC MODEL

Consider the following STAR model of order p

$$y_t = \pi_{10} + \pi_1' w_t + (\pi_{20} + \pi_2' w_t) F(y_{t-d}) + u_t \tag{1}$$

where $u_t \sim \text{nid}(0, \sigma^2)$, $\pi_j = (\pi_{j1}, \ldots, \pi_{jp})'$, $j = 1, 2$, $w_t = (y_{t-1}, \ldots, y_{t-p})'$ and F is a transition function which by convention is bounded by zero and one. We shall consider two different transition functions in (1). The first one is the logistic function

$$F(y_{t-d}) = (1 + \exp[-\gamma(y_{t-d} - c)])^{-1}, \quad \gamma > 0 \tag{2}$$

and the second one is

$$F(y_{t-d}) = 1 - \exp(-\gamma(y_{t-d} - c)^2), \quad \gamma > 0. \tag{3}$$

Writing

$$y_t = (\pi_{10} + \pi_{20}F) + (\pi_1 + \pi_2 F)' w_t + u_t \tag{4}$$

it is clear that (2) allows the 'parameters' in the state-dependent autoregressive model (4) (Priestley 1988) to change monotonically with y_{t-d}. Note that when $\gamma \to \infty$ in (2), $F(y_{t-d})$ becomes a Heaviside function: $F(y_{t-d}) = 0$, $y_{t-d} \leqslant c$, $F(y_{t-d}) = 1$, $y_{t-d} > c$, and (1) with (2) becomes a TAR(p) model. When $\gamma \to 0$, (1) becomes a linear AR(p) model. Model (1) with (2) is called the logistic STAR (LSTAR) model; see e.g. Chan and Tong (1986) or Teräsvirta (1990a,b). Applied to the modelling of business cycle indicators the LSTAR model describes a situation where the contraction and expansion phases of an economy may have rather different dynamics, and a transition (change in dynamics) from one to the other may be smooth. The LSTAR model is also capable of generating asymmetric realizations; for discussion see e.g. Luukkonen and Teräsvirta (1991). The model may be generalized in various ways, for instance by having a linear combination of lags in the exponents in (2) and (3). Only a single lag is used here for reasons of parsimony.

If (1) is accompanied by (3) we have a model in which the 'parameters' in (4) change symmetrically about c with y_{t-d}. Such a model is called an exponential STAR (ESTAR) model; see Teräsvirta (1990a,b). If $\gamma \to \infty$ the ESTAR model becomes linear, but note that this also happens if $\gamma \to 0$, because on the boundary one regime has probability one and the other one probability zero. The ESTAR model implies that the contraction and expansion have rather similar dynamic structures, whereas the middle ground can have different dynamics. An ESTAR model can therefore represent an economy which returns from high growth towards more 'normal' growth in much the same fashion as it accelerates from low or negative growth towards the middle ground. It is obvious that the LSTAR and ESTAR models can describe widely different kinds of dynamic economic behaviour, which makes the STAR family of models a promising family for modelling nonlinearities in business cycles.

When using the STAR family of models in this context there is no economic theory to distinguish between LSTAR and ESTAR models, so that the choice between these models has to be based on the data. Likewise, the delay parameter d in (1), as well as the lag structure of the model, have to be determined from the data. These specification issues are discussed in the next section.

3. SPECIFICATION OF STAR MODELS

As is obvious from (1) and (4), the linear AR(p) model is nested in the STAR model. A natural first step in specifying the model is therefore to test linearity of the model against the STAR specification. If the null of linearity is accepted, the conclusion is that the business cycle indicator can be adequately described by a linear AR model. If it is rejected, the specification of the nonlinear model becomes a matter of concern.

The specification of STAR models is discussed in Teräsvirta (1990a). It consists of three stages:

1. Specification of a linear AR model.
2. Testing linearity for different values of the delay parameter d, and if it is rejected, determining d.
3. Choosing between LSTAR and ESTAR models using a sequence of tests of nested hypotheses.

Stage (1) forms the basis of the linearity testing. Even if the true model is linear, the maximum lag p is usually unknown and has to be determined from the data. A suitable model selection criterion may be used for this purpose. If the true model is nonlinear it is possible that a maximum lag greater than the maximum lag in the nonlinear model will be selected. This may have an adverse effect on the power of the test compared to the case where the maximum lag is known. Saikkonen and Luukkonen (1991) recently discussed this problem in the framework of a simple bilinear model. Another possible problem is that the selected p is so low that the estimated AR model has autocorrelated residuals. In this case the linearity test is then biased towards rejection if the true model is linear, because often the test also has power against serially correlated errors as shown in Teräsvirta (1990a).

From (1), (2) and (3) it is seen that testing H_0: $\gamma = 0$, assuming that y_t is stationary and ergodic under H_0, is a nonstandard testing problem, in the sense that (1) is only identified under the alternative H_1: $\gamma \neq 0$. In this particular case the problem has a simple solution based on an approximation which is outlined in Teräsvirta (1990a), and a Lagrange multiplier (LM) type test of linearity against STAR (both LSTAR and ESTAR) assuming d is known is equivalent to the test of

$$H_0: \beta_{2j} = \beta_{3j} = \beta_{4j} = 0, \; j = 1, \ldots, p$$

against H_1: 'H_0 is not valid' in the artificial (approximating) regression

$$y_t = \beta_0 + \beta_1' w_t + \sum_{j=1}^{p} \beta_{2j} y_{t-j} y_{t-d} + \sum_{j=1}^{p} \beta_{3j} y_{t-j} y_{t-d}^2 + \sum_{j=1}^{p} \beta_{4j} y_{t-j} y_{t-d}^3 + v_t. \tag{5}$$

In order to specify d, the test is carried out for the range of values $1 \leqslant d \leqslant D$ considered appropriate. If linearity is rejected for more than one value of d, then d is determined as $\hat{d} = \arg\min p(d)$ for $1 \leqslant d \leqslant D$ where $p(d)$ is the p-value of the selected test. The argument behind this rule is that the test has maximum power if d is chosen correctly, whereas an incorrect choice of d weakens the power of the test.

After determining d, the purpose of the third stage is to choose between LSTAR and ESTAR. This can be done by a sequence of tests within (5). The sequence of hypotheses to be tested is as follows (Teräsvirta, 1990a):

$$H_{04}: \; \beta_{4j} = 0, \; j = 1, \ldots, p. \tag{6}$$

$$H_{03}: \; \beta_{3j} = 0 \mid \beta_{4j} = 0, \; j = 1, \ldots, p. \tag{7}$$

$$H_{02}: \; \beta_{2j} = 0 \mid \beta_{3j} = \beta_{4j} = 0, \; j = 1, \ldots, p. \tag{8}$$

The logic behind this sequence is based on interpreting the coefficients β_{ij} as functions of the parameters of the original model, i.e. (1) with either (2) or (3). If the model is an ESTAR model, then $\beta_{4j} = 0$, $j = 1, \ldots, p$, but $\beta_{3j} \neq 0$ for at least one j if $\pi_2 \neq 0$. Furthermore, if the model is a LSTAR model, $\beta_{2j} \neq 0$ for at least one j if $\pi_2 \neq 0$. Thus, if (6) is rejected we choose the LSTAR model. If (6) is accepted and (7) is rejected, the ESTAR model is selected. Accepting (6) and (7) and rejecting (8) leads to the LSTAR model. The only inconclusive case is the one in which both (7) and (8) are rejected. In this case we may test

$$H_{03}': \; \beta_{3j} = 0 \mid \beta_{2j} = \beta_{4j} = 0, \; j = 1, \ldots, p.$$

This test is not invariant to the location of y_t. This means that if the test is carried out for a transformed series $y_t^* = y_t + \Delta$, $\Delta \neq 0$, the value of the test statistic changes. We can make use of this fact by performing the test for $y_t + \Delta_j$, $\Delta_1 < \Delta_2 < \cdots < \Delta_k$, $\Delta_1 < 0$, $\Delta_k > 0$, including $\Delta_i = -\bar{y}$. The specification is based on the p-values $p(\Delta_j)$, $j = 1, \ldots, k$. If this p-function has a unique maximum, the LSTAR model is chosen; in the opposite case in which a unique minimum occurs, the ESTAR model is the final choice. See Teräsvirta (1990a) for the motivation of this procedure.

All of this can be done without estimating any nonlinear models. However, to specify the lag structure we suggest the estimation of different nonlinear models using the specified linear AR models as a base, and reducing from general to specific specifications. This part of the search could be formalized by using model selection criteria to guide the search, but assessing the properties of the estimated model should be an important part of the search process. We illustrate this specification technique in the following section of the paper.

4. NONLINEAR MODELLING OF INDUSTRIAL PRODUCTION

4.1. The Data

When studying the possible nonlinearity of business cycles it is useful to choose a business cycle indicator that shows as much cyclical variation as possible. We use international data on the volume of industrial production for our analysis because this series shows more cyclical variation than GNP, which contains some sluggish components. The time series are the same as in Luukkonen and Teräsvirta (1991); the quarterly observations are from the period 1960(i) to 1986(iv) and they are seasonally unadjusted values of the logarithmic indices of industrial production for 13 OECD countries and Europe. The series for Europe is an aggregate of European countries that are OECD members. Observations from 1987(i) to 1988(iv) are used for forecasting. The source is the *OECD Main Economic Indicators*.

The original series are made approximately stationary by seasonal (four-quarter) differencing, and this annual (four-quarter) growth rate series is a reasonable business cycle indicator. The French output series is adjusted for strikes in 1963(i) and 1968(ii), and the Italian series has also been adjusted to eliminate the effects of widespread industrial action in 1970(iv).

4.2. Linearity Tests

The purpose of linearity testing in specification is two-fold. Firstly, we want to find the countries for which linearity is not rejected so that we can exclude them from model building efforts. Secondly, if linearity is rejected, the test provides an estimate for the delay parameter. The test results are in Table I.[1] Note that the p-values relate to tests which have been carried out assuming that the delay parameter d is known. If we wanted to control the overall significance level by assuming that d is unknown, we could modify the linearity test against STAR as in Luukkonen *et al.* (1988) to cover the present situation. However, applying the modified test would require longer series than are presently available, because in the modified

[1] The actual size of the tests may cause concern because the linear AR models are large and the linearity tests require many degrees of freedom. However, we did some simulations to investigate this matter, and found that the F test may be safely applied in this study. An example of our simulations is as follows:

> Using the least squares estimated coefficients for an AR(10) model of four-quarter differences of the European series as parameters, we generated 2000 time-series of 100 observations for an AR(10) model with standard independent normal errors. Initialization effects were avoided by generating 200 observations, and then discarding the first 100 of these observations. We then tested linearity in each of the 2000 generated series using the F statistic based on (5) and a nominal significance level of 0·05. (We set d equal to one for the purposes of the test). There are 30 degrees of freedom in the numerator of the test statistic, but the empirical size of the test is an acceptable 0·041.

This and other experiments we conducted indicated that the F test can be applied in this study without further concern.

Table I. Minimum p-values of linearity tests, chosen values of the delay parameter and the chosen model family based on four quarter differences of industrial production data from 13 OECD countries and Europe 1961(i)–1986(iv)

Country	Maximum lag[a]	Minimum p-value over $1 \leqslant d \leqslant 5$[b]	Corresponding delay (quarters)	Type of model
Austria	5	0·010	1	LSTAR
Belgium	5	0·050	1	LSTAR
Canada	5	0·071	2	LSTAR[c]
FR Germany	9	0·004	4	LSTAR
Finland	1	0·547	—	Linear
France	9	0·156	—	Linear
Italy	5	0·029	3	ESTAR
Japan	5	0·000	1	LSTAR/ESTAR
The Netherlands	1	0·123	—	Linear
Norway	8	0·031	5	LSTAR
Sweden	5	0·016	3	LSTAR
United Kingdom	8	0·047	4	ESTAR
United States	6	0·006	3	LSTAR
Europe	9	0·015	3	LSTAR/ESTAR

[a] The selection was made using AIC.
[b] All of the individual tests (for a fixed d) are F tests whose empirical sizes correspond closely to the nominal size in this application.
[c] Although the minimum p-value is not low, specification of the nonlinear model was attempted, because while H_{04} and H_{03} are clearly accepted with p-values of 0·45 and 0·40 respectively, the p-value of the test of H_{02} equals 0·013.

test the number of degrees of freedom associated with the test statistic is large if the maximum lag of the AR model is large. Here, the overall size of the procedure is not a critical issue, because our aim is to specify a nonlinear model rather than to test a particular theory. If we incorrectly reject the null of linearity and attempt to estimate a nonlinear model, we are likely to discover that a reasonable model which satisfies our criteria cannot be found.

The use of 0·05 as a rather arbitrary threshold p-value leads to the classification of the series for Finland, France and the Netherlands as linear. Specification of a STAR model for Canada is attempted because of other evidence; see the footnote in Table I. The series for the remaining countries and that for aggregated Europe are classified to be nonlinear according to this criterion, and we build nonlinear models for some of these series in the sections which follow.

Table I also reports the selected values of the delay parameter and the choice between the LSTAR and the ESTAR model. The test results leading to these choices have been omitted, but are available from the authors upon request.

4.3. Discussion of the Estimated Models

The specified models are estimated by nonlinear least-squares. Under certain conditions, including stationarity and ergodicity of the series, the estimators of the parameters are consistent and asymptotically normal; see e.g. Tong (1990, Chapter 5). This section presents estimated equations, and in the spirit of Granger and Andersen (1978) we also report the ratio of the residual variance of the nonlinear model to that of the corresponding AR model chosen by AIC. In general we restrict our consideration to nonlinear models for which this ratio is less than 0·9, so as to avoid models that may be spurious. However, we have included discussion on our analysis for Europe, because the estimated LSTAR and ESTAR models for the European series display many properties which characterize the other estimated models. Table III contains some diagnostics associated with these models. Residuals are tested against fourth-order ARCH using the LM test of Engle (1982), and their normality (or presence of outliers) is checked with the Jarque–Bera normality test. The skewness and excess kurtosis of the residuals are also reported.

It is usually difficult to interpret the individual coefficients of STAR models, but the roots of the characteristic polynomials associated with these models provide information which is useful for understanding their dynamic properties. We compute the roots of the STAR(p) model by solving

$$z^p - \sum_{j=1}^{p} (\hat{\pi}_{1j} + \hat{\pi}_{2j}F)z^{p-j} = 0$$

for $F = 0, 1$. It is possible to consider the roots for any value of F (where $0 \leqslant F \leqslant 1$) but the roots of the respective extreme regimes describe the local dynamics of recession and expansion. See Teräsvirta (1990a,b) for further examples.

A useful diagnostic check is to consider the long-run properties of the model

$$y_t = f(y_{t-1}, \ldots, y_{t-p}; \hat{\theta}) \tag{9}$$

where $\hat{\theta} = (\hat{\pi}_{10}, \hat{\pi}_1', \hat{\pi}_{20}, \hat{\pi}_2', \hat{\gamma}, \hat{c})'$. Since we cannot generally solve (9) analytically for y_t, we 'start' (9) with various sets of starting values $y_{t-1}^0, \ldots, y_{t-p}^0$ and observe how the process (with noise suppressed) develops as $t \to \infty$. It may converge to a stable stationary point (Ozaki, 1985), display a limit cycle (i.e. the same set of q values $\bar{y}_t, \bar{y}_{t-1}, \ldots, \bar{y}_{t-q+1}$ which repeats itself over time independent of the starting values), or diverge. In the last case the estimated model is rejected. Another possibility is 'chaotic' behaviour; the process does not diverge, but small changes in starting values can generate quite different solution paths. We see examples of this in the next subsections.

United States

The four-quarter differences, y_t, of the logarithms of US industrial production index are graphed in Figure 1. The most interesting parts of the series are the recessions of 1970, 1974–75, 1980 and 1982, because they are not well explained by the AR(6) model selected using AIC. The specified and estimated LSTAR model for y_t is

$$\begin{aligned} y_t = &\underset{(0\cdot0072)}{-0\cdot021} + \underset{(0\cdot12)}{0\cdot35}y_{t-1} + \underset{(0\cdot20)}{0\cdot24}y_{t-3} - \underset{(0\cdot19)}{1\cdot03}y_{t-4} + \underset{(0\cdot11)}{0\cdot33}y_{t-9} + (\underset{(0\cdot0072)}{0\cdot021} + \underset{(0\cdot15)}{1\cdot16}y_{t-1} - \underset{(0\cdot10)}{0\cdot57}y_{t-2} \\ &- \underset{(0\cdot20)}{0\cdot24}y_{t-3} + \underset{(0\cdot19)}{1\cdot03}y_{t-4} - \underset{(0\cdot11)}{0\cdot33}y_{t-9}) \times (1 + \exp[-\underset{(37)}{49} \times 17\cdot5(y_{t-3} - \underset{(0\cdot0007)}{0\cdot0061})])^{-1} + \hat{u}_t \end{aligned} \tag{10}$$

$$s = 0\cdot0176, \quad s^2/s_L^2 = 0\cdot64.$$

where s is the residual standard deviation of (10), s_L is the corresponding statistic for the AR(6) model, and the figures in parentheses are the estimated standard errors. It is seen that the error variance of (10) is considerably less than that of the AR(6) model. The restrictions $\pi_{1j} = -\pi_{2j}$, $j = 0, 3, 4, 9$, are suggested by the data and are imposed. Note that while the restriction $\pi_{1j} = 0$ makes the coefficient of lag j zero when the transition function F is zero, the restriction $\pi_{1j} = -\pi_{2j}$ does the same when the transition function F is one. Both restrictions are therefore exclusion restrictions. The long maximum lag and the gap between lags 4 and 9 are due to the

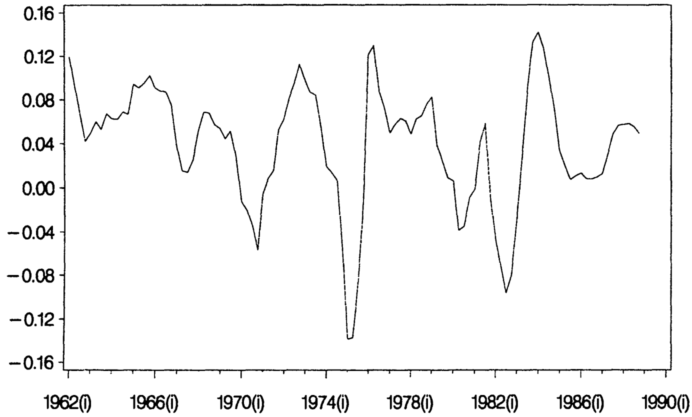

Figure 1. Four-quarter differences of the quarterly logarithmic index of US industrial production, 1962(i) to 1988(iv)

fact that the series is not seasonally adjusted. Moreover, the model does not contain any moving-average terms. The interpretable parameter estimates in (10) are $\hat{\gamma} = 49/\hat{\sigma}(y)$ and $\hat{c} = 0{\cdot}0061$. For the latter, $\hat{F} = \frac{1}{2}$, so that $\hat{c}$ marks the halfway point between the recession and the expansion, and it is close to zero at 0·6 per cent. (Note that γ has been scaled by $\hat{\sigma}(y) = 1/17{\cdot}5$ as Teräsvirta (1990a) suggested.) The value of $\hat{\gamma} = 49/\hat{\sigma}(y)$ indicates that the transition from one regime to the other is very quick so that the model is very similar to a TAR model. The standard deviation of $\hat{\gamma}$ is relatively large and is an indication of the fact that accurate estimation is a problem when y_{t-3} is close to c and F increases rapidly. A broad range of values of γ will then give about the same F, and many observations in the neighbourhood of c are needed to obtain a precise estimate of γ. This situation is described in detail in Bates and Watts (1988, p. 87). The delay parameter is estimated to be $\hat{d} = 3$. This, combined with $\hat{c}$ close to zero, suggests that the regimes change when the *level* of the series reaches a peak or a trough. When the realization of a growth rate process appears smoothly cyclical, the annual growth rate crosses zero from above some quarters after the level has peaked, and hence a delay which is longer than one quarter is observed. An analogous lag is present at the trough

The dynamics of the model can be deduced from Table II. The most prominent pair of complex roots in the lower or recession regime has a modulus of 1·1 and a period of 8·9 quarters, so that the process is locally explosive. On the other hand, the upper or expansionary regime is completely characterized by a complex pair of roots which have modulus 0·76 and a period of 61 quarters. This asymmetry of regimes is the most striking feature of the model. The production moves from deep recession into higher growth very aggressively, whereas there is nothing in the dynamics of the expansionary regime to suggest a rapid fall into a contraction. Only a sufficiently large negative shock could cause this.

Figure 2 is useful in assessing the benefits of the LSTAR model. It is seen that the aftermath of both the first and second oil crises is much better explained by (10) than by the AR(6) model. The same is true for the recession of 1970. The AR model is not able to foresee a rapid return to positive growth. This failure was particularly dramatic after the first oil shock and still visible after the second shock. On the other hand, if we assume that the two negative oil shocks were exogenous to the system we might expect large negative residuals in both (10) and the AR(6) corresponding to these events, as Figure 1 shows. Note also the negative residuals in 1970(i) and 1970(ii).

Table II. Limiting behaviour ($t \to \infty$) of the estimated models and the most prominent roots of the characteristic polynomials in both regimes

Country	Limiting behaviour[a]	Most prominent roots Regime[b]	Root(s)	Modulus	Period
United States	Chaotic behaviour;	L	0·85 ± 0·72i	1·11	8·9
	long and short cycles;	U	0·75 ± 0·08i	0·76	61·0
	min ~ −0·06				
	max ~ 0·04				
Europe	USSP (0)	L	0·55 ± 0·85i	1·01	6·3
		U	0·95	0·95	
			0·80 ± 0·39i	0·89	13·9
Belgium	USSP (0·039)	L	0·77 ± 0·68i	1·02	8·6
		U	0·61 ± 0·58i	0·83	8·3
Canada	Chaotic behaviour;	L	0·82 ± 0·62i	1·03	9·7
	long and short cycles;	U	0·57 ± 0·51i	0·76	8·7
	min ~ −0·06				
	max ~ 0·07				
FR Germany	USSP (0)	L	0·81 ± 0·65i	1·04	9·2
		U	0·84 ± 0·31i	0·90	18·2
		U	0·87	0·87	
Italy	USSP (0·055)	M	1·03	1·03	
		M	−0·55	0·55	
		O	0·85 ± 0·53i	1·00	11·4
Japan	(i) SSP (0·080)				
	(ii) Chaotic behaviour;	M	2·48	2·48	
	cycles with length	O	0·94	0·94	
	11–13 quarters,	O	0·81 ± 0·48i	0·94	11.7
	min ~ −0·21				
	max ~ 0·29				

[a] USSP = Unique stable singular point, SSP = Stable singular point.
[b] L = Lower regime, U = Upper regime, M = Mid-regime, O = Outer regime.

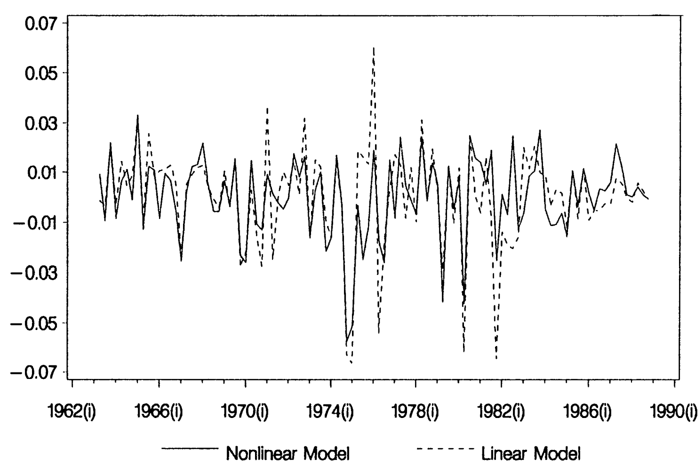

Figure 2. Residuals of (10) and an AR(6) model of the four-quarter differences of the quarterly logarithmic index of US industrial production, 1963(ii) to 1988(iv)

Table III. p-values of fourth-order ARCH and Jarque–Bera normality tests, and skewness and excess kurtosis measures of residuals from the estimated nonlinear models

	p-values		Other statistics	
Country	ARCH(4) test	Jarque–Bera test	Skewness	Excess kurtosis
Belgium	0·39	0·86	−0·14	0·05
Canada	0·70	0·93	−0·07	0·15
FR Germany	0·27	0·63	−0·19	0·30
Italy	0·36	0·03	−0·66	0·45
Japan	0·18	0·68	−0·17	−0·31
United States	0·41	0·0006	−0·85	1·15
Europe	0·36	0·45	−0·17	0·58

The diagnostic statistics are presented in Table III. There is evidence of negative skewness and positive excess kurtosis, which seems to be due to the residuals corresponding to the two oil shocks. The STAR model does not explain these shocks because they are exogenous.

Model (10) offers an interesting description of the dynamics of the US industrial output over 1961(i)–1986(iv), particularly with respect to long-run behaviour. As Table II indicates, this model generates both short and long cycles whose characteristics depend on starting values. Figure 3 illustrates two examples of this 'chaotic' behaviour: both sets of observations have been generated from (10) without noise; the starting values for the first series are the nine four-quarter differences of the volume series preceding 1980(i), while the second set of starting values are the same, excepting that the 1979(iv) starting value has been perturbed by 0·001. After 100 observations the solution paths are quite different from each other. Such behaviour

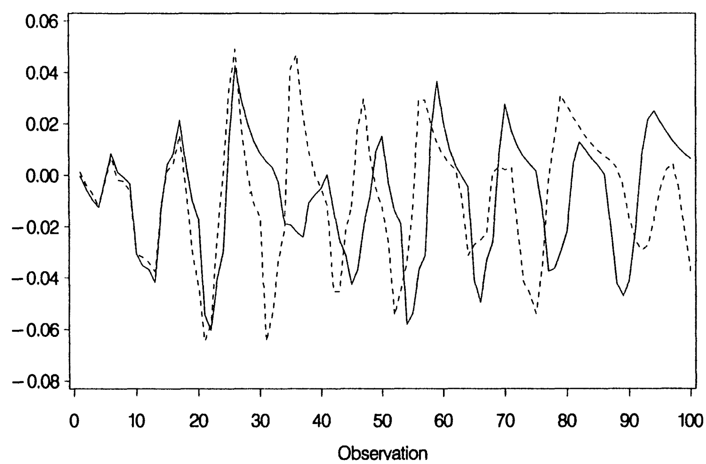

Figure 3. Two realizations generated from model (10) with the noise suppressed. The starting values for the series depicted by the joined line are the nine four-quarter differences of the logarithmic index of US industrial production preceding 1980(i). The starting values for the series depicted by the dashed line are the same, except that the 1979(iv) value has been perturbed by 0·001. The graph shows the first 100 observations generated in each case

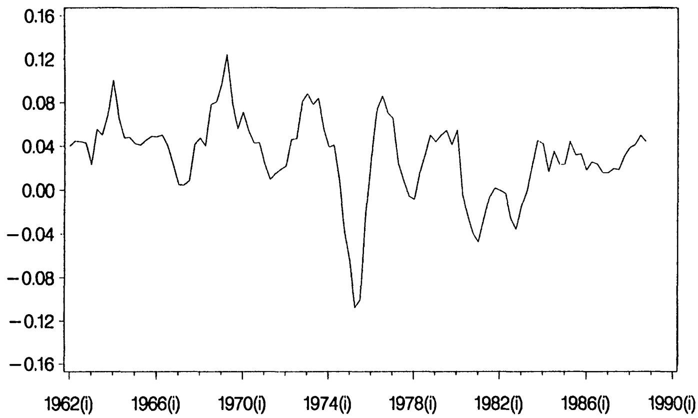

Figure 4. Four-quarter differences of the quarterly logarithmic index of European industrial production, 1962(i) to 1988(iv)

illustrates that the long-run effect of any shock on the growth rate of industrial production depends on the previous history of the process.

Europe

We consider European industrial output as an aggregate before modelling the European countries individually. The graph of the series appears in Figure 4. The dominating feature is the large drop in the growth rate as a result of the first oil crisis; the second period of interest is the years 1980–1982. As Table I shows, the linearity tests suggest nonlinearity, but the choice between LSTAR and ESTAR models is not clear-cut (results are not shown). Both models were estimated and fitted the data equally well. Here we report the LSTAR model, whose estimated equation is

$$\begin{aligned} y_t = {} & \underset{(0\cdot28)}{0\cdot44}y_{t-1} - \underset{(0\cdot26)}{1\cdot17}y_{t-4} + \underset{(0\cdot13)}{0\cdot56}y_{t-5} - \underset{(0\cdot38)}{0\cdot89}y_{t-8} + \underset{(0\cdot10)}{0\cdot26}y_{t-9} + (\underset{(0\cdot27)}{0\cdot69}y_{t-1} \\ & + \underset{(0\cdot28)}{0\cdot47}y_{t-4} + \underset{(0\cdot38)}{0\cdot59}y_{t-8}) \times (1 + \exp[-5\cdot4 \times \underset{(8\cdot3)}{59\cdot5}(y_{t-3} + \underset{(0\cdot016)}{0\cdot022})])^{-1} + \hat{u}_t \end{aligned} \tag{11}$$

$$s = 0\cdot0165, \quad s^2/s_L^2 = 0\cdot93$$

Note the joint uncertainty in the estimation of γ and c. The diagnostics in Table III do not reveal anything unusual. Table II shows that the lower regime is explosive, albeit very mildly so, whereas the upper regime is stationary. The long-run solution of (11) has a single stationary point which equals zero. Thus in contrast to (10), which predicts that the long-run effect is dependent on the starting value, a shock to the system which affects starting values does not have a long-run effect on the growth rate.

The residuals of (11) and those of the corresponding AR(9) model are shown in Figure 5. The figure shows clearly that the main need for a nonlinear model arises from the first oil crisis. Elsewhere the fits of the linear and LSTAR model agree closely, and the gain from the nonlinear model fit measured by the ratio of residual variances remains rather small. The inevitable conclusion is that the LSTAR model describes the recovery of output from the first oil shock better than the linear model.

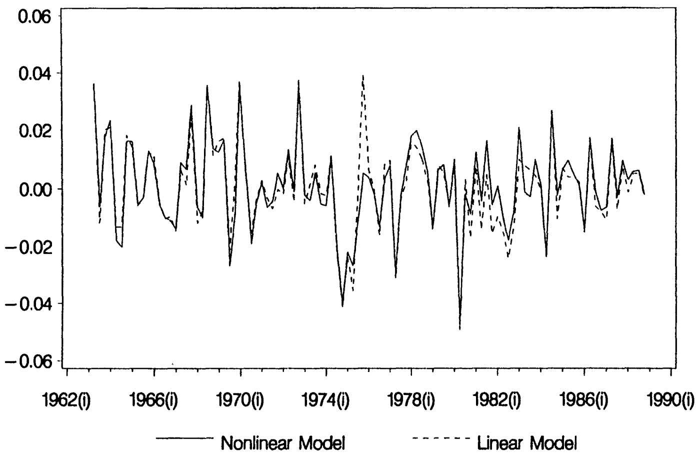

Figure 5. Residuals of (11) and an AR(9) model of the four-quarter differences of the quarterly logarithmic index of European industrial production, 1963(ii) to 1988(iv)

Japan

Linearity is rejected very strongly for the Japanese industrial production series. A major reason for this is a very deep trough of $-0{\cdot}19$ during the first oil crisis and the subsequent rapid recovery. As is seen in Figure 6, this is the dominant feature of the Japanese series. The choice between ESTAR and LSTAR models is again difficult. The estimated ESTAR model is

$$
\begin{aligned}
y_t = \underset{(0\cdot0034)}{0\cdot0075} + \underset{(0\cdot40)}{3\cdot03}y_{t-1} - \underset{(0\cdot19)}{1\cdot31}y_{t-2} - \underset{(0\cdot089)}{0\cdot49}\Delta y_{t-4} + (-\underset{(0\cdot39)}{1\cdot68}y_{t-1} + \underset{(0\cdot24)}{0\cdot87}y_{t-2} \\
- \underset{(0\cdot087)}{0\cdot30}\Delta y_{t-8}) \times [1 - \exp(-\underset{(0\cdot88)}{1\cdot54} \times 196(y_{t-1} + \underset{(0\cdot012)}{0\cdot082}))^2] + \hat{u}_t \quad (12)
\end{aligned}
$$

$$s = 0\cdot0185, \quad s^2/s_L^2 = 0\cdot78.$$

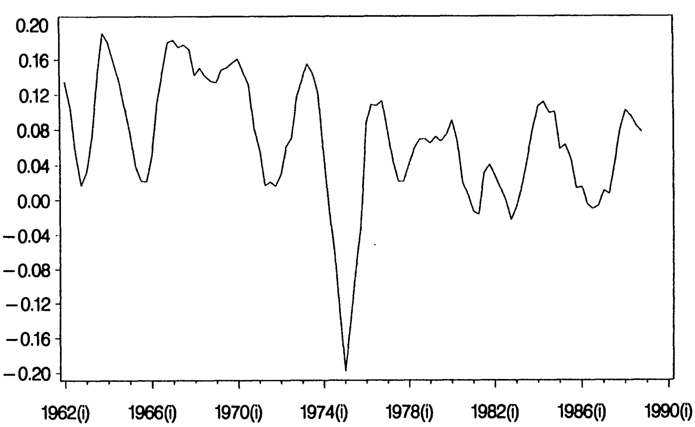

Figure 6. Four-quarter differences of the quarterly logarithmic index of Japanese industrial production, 1962(i) to 1988(iv)

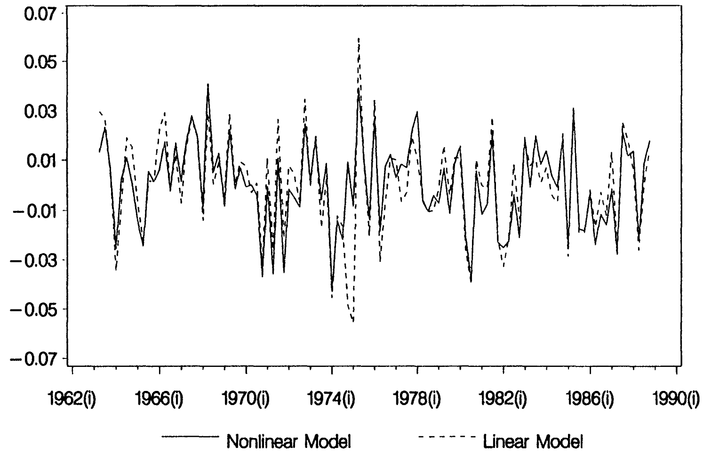

Figure 7. Residuals of (12) and an AR(5) model of the four-quarter differences of the quarterly logarithmic index of Japanese industrial production, 1963(ii) to 1988(iv)

A characteristic feature of (12) is that $\hat{c} = -0{\cdot}082$ is low, so that most observations exceed this value and fall in the right-hand tail of the exponential function. The model therefore performs like an LSTAR model, except for 1974–75 when the left-hand tail is in use. Table II shows that the mid-regime $y_{t-1} = -0{\cdot}082$ contains a single explosive root, so that the process passes this value quickly on the way up or down. The outer regime is stationary, and the most prominent pair of roots has a 3-year period.

The residuals in Figure 7 tell a similar story to that for the LSTAR model of European industrial production. The nonlinearity is there to explain the rapid recovery from the oil shock in 1974–75. The linear model and (12) track industrial production equally well during the second oil crisis in 1979–80, so that the country absorbed lessons from the first crisis very rapidly. In general we cannot conclude that the series is inherently nonlinear. The diagnostics in Table III do not indicate any model inadequacy.

The numerical long-run solution depends on the starting values. If the starting values are positive, the series converges to a stationary point corresponding to an annual growth rate of 8 per cent. If the starting values are taken from a period which includes the deep trough, the generated series contain 11–13 quarter cycles with wide oscillations, and any slight change in these starting values will cause substantial differences in the observed solution paths.

It may be worth mentioning that the ESTAR model which was fitted to the European series but not reported in the previous section resembles (12) in that $\hat{c} = -0{\cdot}11$ is large in absolute value and negative, and the mid-regime is explosive whereas the outer one is stationary. This is not surprising, because the ESTAR model fits the European data as well as the LSTAR model (11), and behaves like an LSTAR model. Its characterization of the European series is similar to that of (11) in many ways, just as the ESTAR model (12) for Japan also behaves like an LSTAR model.

Belgium, Canada and FR Germany

These three countries are studied together because the nonlinear models for their industrial production are all LSTAR models with interesting common features. The estimated model for

Belgium is

$$\underset{(0\cdot018)}{y_t = -0\cdot030} + \underset{(0\cdot11)}{0\cdot64y_{t-1}} - \underset{(0\cdot20)}{0\cdot29y_{t-2}} - \underset{(0\cdot088)}{0\cdot64y_{t-4}} + (\underset{(0\cdot021)}{0\cdot044} + \underset{(0\cdot20)}{0\cdot49y_{t-2}}$$

$$+ \underset{(0\cdot082)}{0\cdot45y_{t-5}}) \times [1 + \exp(-\underset{(5\cdot3)}{7\cdot3} \times 21\cdot6(y_{t-1} + \underset{(0\cdot0060)}{0\cdot015}))]^{-1} + \hat{u}_t \quad (13)$$

$$s = 0\cdot0231, \quad s^2/s_L^2 = 0\cdot84.$$

The corresponding LSTAR model for Canadian industrial production is

$$y_t = \underset{(0\cdot064)}{1\cdot29y_{t-1}} - \underset{(0\cdot12)}{0\cdot64y_{t-3}} - \underset{(0\cdot12)}{0\cdot35y_{t-4}} + \underset{(0\cdot089)}{0\cdot55y_{t-5}} + (\underset{(0\cdot014)}{0\cdot027} - \underset{(0\cdot15)}{0\cdot56y_{t-1}}$$

$$+ \underset{(0\cdot12)}{0\cdot64y_{t-3}} - \underset{(0\cdot10)}{0\cdot34y_{t-5}}) \times [1 + \exp(-\underset{(50)}{37} \times 19\cdot8(y_{t-2} - \underset{(0\cdot0042)}{0\cdot060}))]^{-1} + \hat{u}_t \quad (14)$$

$$s = 0\cdot0179, \quad s^2/s_L^2 = 0\cdot80.$$

Successful estimation required rescaling γ. The estimated LSTAR model for West Germany is taken from Teräsvirta (1990a) and has the form

$$y_t = \underset{(0\cdot070)}{0\cdot82y_{t-1}} - \underset{(0\cdot087)}{0\cdot47y_{t-4}} + \underset{(0\cdot14)}{0\cdot43y_{t-7}} + (\underset{(0\cdot10)}{0\cdot35y_{t-2}} + \underset{(0\cdot12)}{0\cdot24y_{t-5}} - \underset{(0\cdot15)}{0\cdot43y_{t-7}} - \underset{(0\cdot12)}{0\cdot36y_{t-8}}$$

$$+ \underset{(0\cdot11)}{0\cdot30y_{t-9}}) \times [1 + \exp(-\underset{(49000)}{1600} \times 33\cdot5(y_{t-4} - \underset{(0\cdot0070)}{0\cdot0070}))]^{-1} + \hat{u}_t \quad (15)$$

$$s = 0\cdot0250, \quad s^2/s_L^2 = 0\cdot89.$$

Both (14) and (15) are representative examples of the situation where γ is large and estimating it accurately is difficult, as discussed in Bates and Watts (1988, p. 87) and Teräsvirta (1990a). The threshold value of $\hat{c} = 0\cdot060$ for Canada is considerably higher than those for Belgium or West Germany.

Table III shows that each of the models passed the usual diagnostic checks. Table II demonstrates the fact that the characteristic polynomials of the lower regimes in each of the three models have an explosive pair of roots, whereas the upper regimes are stationary. The Belgian and West German models have a unique stable stationary point, but the model for Canadian industrial output generates chaotic behaviour, as did the model for the US. All three models suggest possible asymmetry of business cycles in the sense what the dynamics of the recession are different from the dynamics of the expansion.

Italy

The remaining country whose industrial production index we have been able to model using the STAR family is Italy. The Italian industrial production series is characterized by sharp peaks or troughs, as illustrated by Figure 8, so that the ESTAR model may offer a plausible parameterization of such behaviour. Table I shows that the data indeed suggest an ESTAR model for Italy. Its estimated equation is

$$y_t = \underset{(0\cdot18)}{0\cdot48y_{t-1}} + \underset{(0\cdot20)}{0\cdot57y_{t-2}} + (\underset{(0\cdot0079)}{0\cdot0090} + \underset{(0\cdot28)}{0\cdot91y_{t-1}} - \underset{(0\cdot30)}{0\cdot98y_{t-2}} - \underset{(0\cdot24)}{1\cdot01y_{t-4}} + \underset{(0\cdot27)}{0\cdot86y_{t-5}}$$

$$- \underset{(0\cdot18)}{0\cdot45y_{t-8}} + (\underset{(0\cdot15)}{0\cdot31y_{t-9}}) \times [1 - \exp(-\underset{(0\cdot56)}{1\cdot25} \times 273(y_{t-3} - \underset{(0\cdot0067)}{0\cdot028})^2)] + \hat{u}_t \quad (16)$$

$$s = 0\cdot0279, \quad s^2/s_L^2 = 0\cdot81.$$

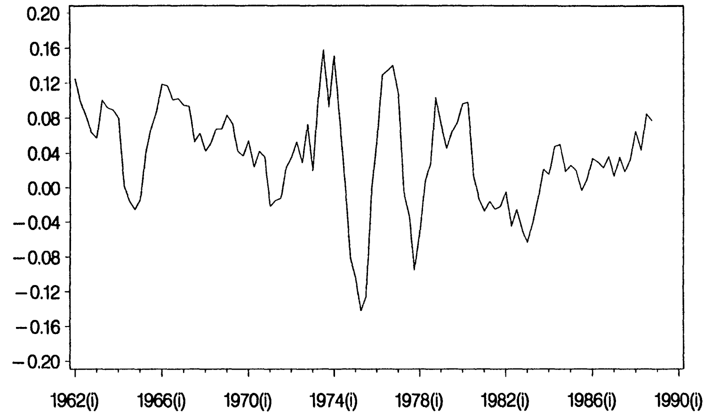

Figure 8. Four-quarter differences of the quarterly logarithmic index of Italian industrial production, 1962(i) to 1988(iv)

Equation (16) is a typical ESTAR model; $\hat{c} = 0 \cdot 028$, so that there are many observations in both tails of the exponential function. Table II shows that there is an explosive complex pair of roots associated with the outer regime, and this makes the process ultimately swing towards $\hat{c}$. On the other hand, the mid-regime has a single explosive root so that the growth rate will pass $\hat{c}$ quickly on the way up or down. Nevertheless, the model as a whole has a unique stationary point. The diagnostics in Table III find negative skewness in the residuals. This may be due to exogenous negative shocks, as discussed previously.

5. FORECASTING WITH STAR MODELS

Another way of evaluating the estimated nonlinear models is post-sample forecasting, although the insight to be gained depends on what happens in the time-series during the prediction period. Generally a wide range of values over the prediction period is needed to efficiently compare the forecasting performance of a STAR model to that of a linear autoregressive model. Our previous results suggest that if the prediction period does not contain a clear recession, then the linear and nonlinear forecasts will be similar, unless the specification of the STAR model is grossly inadequate.

The prediction period used for our forecasting (1987 and 1988) was indeed rather 'normal' for the countries (and Europe) for which we were able to specify and estimate STAR models of industrial production. The results of our forecasts are presented in Table IV. This table contains the root mean square errors (RMSE) of the eight quarterly one-step-ahead forecasts for the linear and nonlinear models of each series. The forecasts were made without re-estimating the models during the prediction period. In all cases the RMSE of both the AR and STAR models lie below the residual standard deviation of the model, indicating that the period was uneventful and therefore relatively easy to predict. The US industrial output provides the best illustration of this observation; the RMSE of the linear forecasts is very small, only 0·4 per cent and that of the LSTAR forecasts is only one half of the residual standard deviation of (10).

Table IV also contains results of testing the hypothesis that there is no difference in the prediction accuracy between the linear and non-linear models. We applied the test in Granger and Newbold (1986, pp. 278–280) which is based on the correlation coefficient of the sums and

Table IV. Root mean squared errors of the forecasts for the linear and nonlinear models and p-values of the test H_0: $r = 0$ vs H_1: $r > 0$

Country	Root mean squared errors		p-Value of the test
	Linear model	Nonlinear model	H_0: $r = 0$
United States	0·0041	0·0092	0·995
Europe	0·0082	0·0080	0·43
Japan	0·0184	0·0175	0·39
Belgium	0·0165	0·0158	0·37
Canada	0·0146	0·0181	0·85
FR Germany	0·0201	0·0176	0·26
Italy	0·0187	0·0184	0·48

differences of the forecast errors. The null hypothesis is H_0: $r = 0$, and the test is carried out assuming the one-sided alternative that $r > 0$. Our alternative is that the nonlinear model has superior forecasting performance. The null distribution is based on the well-known approximation that

$$w = \tfrac{1}{2}\ln\left(\frac{(1+r)}{(1-r)}\right) \sim N\left(0, \left(\frac{1}{N-3}\right)\right)$$

where N is the number of forecasts ($N = 8$). As expected, the null hypothesis was not rejected in any of the cases considered. The results for the US suggest the superiority of the AR forecasts, but as pointed out above, both forecasts are very accurate. Only one out of the eight forecast errors from the LSTAR model (10) exceeds the residual standard deviation, and this error is the cause of the high p value found in Table IV. As the LSTAR model yields remarkably accurate forecasts for 1987–88, it can hardly be rejected on the basis of its forecasting performance relative to that of a linear model. We conclude that our prediction period does not give us any new information either against or in favour of the seven estimated STAR models.

6. CONCLUSION

The estimated STAR models for industrial growth in different OECD countries have similar features. Most of the models suggest that the dynamics of the process during recession are different from those during expansion. The characteristic polynomial of the recession regime usually displays at least one complex pair of explosive roots while the expansionary regime is stationary. This suggests that business cycles are asymmetric in the sense that the recovery from a deep recession may be strong, swift, and well described by the nonlinear models used, whereas there is no corresponding mechanism leading back to a recession. This seems particularly true for the US business cycle. The estimated models suggest that deep contractions are caused by exogenous shocks, and they also seem to indicate that special measures are taken during recession that generate a speedy recovery of industrial output.

All of these results are, of course, rather time-specific, because the two dominating features during the relatively short observation period are the exogenous shocks associated with the first and second oil crises. In some cases, notably Europe and Japan, the nonlinearity is needed just to model the response to a single large negative shock. We can hardly argue that business cycles in these economies are inherently nonlinear on this basis, but the response to large negative shocks is stronger than that predicted by a linear autoregressive model.

The fact that the response of industrial production to exogenous shocks is similar for many countries, including all of the large economies (excepting France), lends credibility to our modelling exercise. The Italian industrial production is a special case because its growth rate has fluctuated widely both upwards and downward. The STAR models do not seem to be appropriate for modelling four of the series. This may be due to the absence of nonlinearities or to the nature of the nonlinearities present in the data. Alternatively, it is possible that linearity is rejected because of outliers, and STAR modelling of the series is therefore not appropriate. However, the general conclusion is that in many cases the STAR model family is a suitable tool for describing the nonlinearity found in OECD industrial output series.

ACKNOWLEDGEMENTS

This paper was prepared for the Conference on Nonlinear Dynamics and Econometrics organized by the *Journal of Applied Econometrics* and the UCLA Program in Applied Econometrics, 5–6 April 1991. It was also presented at the Econometric Society European Meeting, Cambridge, England, September 1991. The comments from the participants of these meetings, as well as those in seminars at Aarhus, Copenhagen, Göteborg, Kiel, Lund, Oulu, Umeå, Uppsala, Stockholm School of Economics, Southern Methodist University, University of California, San Diego and the Bank of Norway, are greatly appreciated. We are also grateful to Clive Granger for several helpful discussions and Shakir Hussain, Henry Ohlsson, Simon Potter and the anonymous referees for useful suggestions and comments. The responsibility for any errors and shortcomings remains ours. The work was done while the first author was visiting the Department of Economics, University of California, San Diego, and the hospitality of the department during his stay is gratefully acknowledged. His research was supported by the Yrjö Jahnsson Foundation.

REFERENCES

Bates, D. M., and D. G. Watts (1988), *Nonlinear Regression Analysis and its Applications*, Wiley, New York.

Chan, K. S., and H. Tong (1986), 'On estimating thresholds in autoregressive models', *Journal of Time Series Analysis*, **7**, 179–190.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation', *Econometrica*, **50**, 987–1007.

Granger, C. W. J., and A. P. Andersen (1978), *An Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Göttingen.

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

Keynes, J. M. (1936), *The General Theory of Employment, Interest and Money*, Macmillan, London.

Luukkonen, R., and T. Teräsvirta (1991), 'Testing linearity of economic time series against cyclical asymmetry', *Annales d'économie et de statistique*, **20/21**, 125–142.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988), 'Testing linearity against smooth transition autoregressive models', *Biometrika*, **75**, 491–499.

Mitchell, W. C. (1927), *Business Cycles. The Problem and its Setting*, National Bureau of Economic Research, New York.

Ozaki, T. (1985), 'Non-linear time series models and dynamical systems', in E. J. Hannan, P. R. Krishnaiah and M. M. Rao (eds), *Handbook of Statistics*, Vol. 5, Elsevier, Amsterdam, pp. 25–83.

Priestley, M. B. (1988), *Non-linear and Non-stationary Time Series Analysis*, Academic Press, London and San Diego.

Saikkonen, P., and R. Luukkonen (1991), 'Power of a time series linearity test against a bilinear alternative', *Statistica Sinica*, **1**, 453–464.

Teräsvirta, T. (1990a), 'Specification, estimation, and evaluation of smooth transition autoregressive models', Department of Economics, University of California, San Diego, Discussion Paper No 90-39, revised version.

Teräsvirta, T. (1990b), 'Generalizing threshold autoregressive models', Department of Economics, University of California, San Diego, Discussion Paper No 90-44.

Tiao, G. C., and R. S. Tsay (1991), 'Some advances in nonlinear and adaptive modeling in time series analysis', University of Chicago Graduate School of Business and Statistics Research Center Technical Report No 118.

Tong, H. (1990), *Non-linear Time Series. A Dynamical System Approach*, Oxford University Press, Oxford.