Empirical Economics (2002) 27:569–586

EMPIRICAL ECONOMICS

Conditional mean functions of non-linear models of US output

Michael P. Clements[1] **and Ana B. C. Galvão***[2]

[1] Department of Economics, University of Warwick, Coventry, CV4 7AL, UK
(email: M.P.Clements@Warwick.ac.uk)
[2] Department of Economics, European University Institute, Italy.
(email: agalvao@iue.it)

First version received: Feb. 1999/Final version received: June 2001

Abstract. We compare a number of models of post War US output growth in terms of the degree and pattern of non-linearity they impart to the conditional mean, where we condition on either the previous period's growth rate, or the previous two periods' growth rates. The conditional means are estimated non-parametrically using a nearest-neighbour technique on data simulated from the models. In this way, we condense the complex, dynamic, responses that may be present in to graphical displays of the implied conditional mean.

Key words: non-linearity, business cycles, non-parametric conditional mean estimates.

JEL classification: C14, C22

1. Introduction

The last decade has seen the advent of a large number of non-linear time series models to explain the business cycle characteristics of US GNP. These models can generally be given a regime-switching interpretation. The simplest conceptualisation is of an economy characterised by two regimes, say, expansion and contraction, which are modelled as distinct linear autoregressive processes, along with a mechanism governing the movement of the economy between the

* We are grateful to Dick van Dijk, Adrian Pagan and two anonymous referees for helpful comments. Financial support from the UK Economic and Social Research Council under grant L116251015 is gratefully acknowledged by the first author, and from Capes-Brazil and UNISINOS by the second author. The computations reported in this paper were performed using code written in the Gauss Programming Language. We are grateful to Don Harding for providing the code to simulate the Hamilton (1989) and Durland and McCurdy (1994) models.

regimes. The models include the Markov-switching autoregressive model (MS-AR) of Hamilton (1989) and various extensions (for example, to more states, allowing duration dependence in the transition probabilities); threshold and smooth transition autoregressive models, such as Tiao and Tsay (1994), Potter (1995); and models with a continuum of 'regimes', such as the current depth of recession (CDR) models (Beaudry and Koop, 1993) and 'floor and ceiling' models (Pesaran and Potter, 1997).

Whilst traditional econometric approaches to model evaluation emphasize the fit of the model to the sample path of the actual data, foremost in the evaluation of the models just described has often been an assessment of their ability to characterize certain features of the business cycle, such as their ability to replicate the timings of turning points and the durations of expansions and contractions observed in the data, based, for example, on the NBER business cycle chronology of peaks and troughs.[1]

The non-linear time series models are of course subjected to testing, particularly against linearity, to see whether the non-linear structure is necessary, but for certain of these models, this is complicated by the so-called 'Davies problem' (Davies, 1977, 1987), whereby nuisance parameters are unidentified under the null hypothesis of linearity, and standard tests with conventional critical values are invalid.[2] Moreover, as noted by Pagan (1999b), tests may sometimes reject linearity in favour of the non-linear model because of a 'few influential points in the data'. The implication is that for the majority of the data points the non-linear parts of the model are largely irrelevant and are picking up outliers. Testing of one non-linear model against another is rarely done, and such comparisons as there are between non-linear models usually rest on out-of-sample assessments of forecast performance. However, empirical evaluations of the out-of-sample performance of one non-linear model against another are often based on small numbers of forecasts and may be very dependent on the forecast period used, just as the dependence of comparisons between non-linear and linear models on the state of the economy has been documented by a number of authors (e.g., Tiao and Tsay, 1994, Clements and Smith, 1999).

Impulse response analysis is often applied to multivariate linear autoregressive models, i.e., vector autoregressions (VARs), to summarise their often complex dynamics. Thus, the impact of a shock through time can be traced out. Similar types of analyses can be applied to non-linear models: e.g., Koop, Pesaran and Potter (1996) develop generalized non-linear impulse response functions (GIs) (see also Gallant, Rossi and Tauchen, 1993) to analyse the response of non-linear models to shocks. The application of impulse response analysis to non-linear models is complicated by the fact that the impact of the shock is dependent upon the sign and size of the shock, and the 'state' the process is

[1] Pagan (1999a) and Harding and Pagan (2001) have criticised the comparison of the properties or features of the states calculated by models such as the MS-AR model, with the features of the NBER-dated phases, on the grounds that the MS-AR model states are not in principle the same as the NBER classical cycle phases. Nevertheless, comparisons of this sort are widespread in the literature.

[2] Hansen (1996b) and Hansen (1992, 1996a) have proposed valid tests for threshold models and MS-AR models respectively, and for models where the movement between regimes is 'smooth', Taylor approximations can be used to generate test statistics with standard dstributions (see, e.g., Luukkonen, Saikkonen and Teräsvirta, 1988).

in when the shock hits – tracing out the effects through time of a single normalised shock is inappropriate. Consequently, Koop *et al.* (1996) consider the distribution of responses using GIs. A recent paper by van Dijk, Franses and Boswijk (2000) extends the application of impulse response analysis to non-linear models by considering ways of measuring the speeds of absorption of shocks, rather than just their persistence or magnitude.

GIs and related approaches are a useful way of shedding some light on the often complex dynamics inherent in non-linear models, a point stressed by van Dijk *et al.* (2000). We adopt an alternative approach in this paper, which has the same aim. We compare the models in terms of the degree and pattern of non-linearity they impart to the conditional mean, $\mathsf{E}[y_t \,|\, y_{t-1}, \ldots y_{t-p}]$, where $p = 1$ or 2, to allow plotting of the conditional mean function (or surface). Thus we condense the complex, dynamic, non-linearities that may be present in the model in to one or two plots of the implied conditional mean. These are estimated non-parametrically (as explained in section 3), and show how the model prediction depends on the present (and past) growth rate(s). This is a useful summary measure, considered by Pagan (1999b) with $p = 1$, only, and applied to the Pesaran and Potter (1997) 'floor and ceiling', the Beaudry and Koop (1993) CDR, the Potter (1995) SETAR and a three-regime MS-AR model (as given in Hess and Iwata, 1997), and applied by Harding and Pagan (2001) to the Hamilton (1989) two-regime MS-AR and the Durland and McCurdy (1994) MS-AR with duration dependence (again with $p = 1$ only).

So our method of comparing the models is to estimate conditional mean functions on data simulated from the estimated models. We view this approach as being complementary to the calculation of GIs, and to the more common approach of examining whether certain business cycle features of the simulated data from the estimated models matches up with the features we observe in the actual data: see, e.g., King and Plosser (1994), Watson (1994), Hess and Iwata (1997), Pagan (1997b, 1997a, 1999a), Harding and Pagan (2001), Clements and Krolzig (1999) and Galvão (2000). The conditional mean functions are more of a 'broad-brush' approach, and do not allowed a detailed examination of the models' ability to reproduce specific features of the business cycle, but by way of compensation, do not have the attendant difficulties of deciding how to date turning points to define the expansion and contraction phases, or of how to choose the features to match.

In section 2 the salient features of the models are described, to facilitate the discussion in section 4 of how these give rise to the different types and degrees of non-linearity in the conditional mean functions. Section 5 concludes.

2. Models

In line with the recent literature, the linear model is a first-order autoregression in (100 times) the difference of the log of GNP, denoted y_t, which estimated for the period 1947:3 to 1997:1 yields:

$$y_t = 0.5299 + 0.3428 y_{t-1} + \varepsilon_t$$

with the estimated standard error, $\sigma_\varepsilon = 1.007$, and ε_t assumed to be iid. Then, of course, $\mathsf{E}[y_t \,|\, Y_{t-1} = y_{t-1}] = \alpha y_{t-1}$ where $\alpha = 0.3428$, and α does not depend on Y_{t-1}. The conditional expectations for the non-linear models depend on the past of the process.

2.1. Simple threshold models

Threshold autoregressive (TAR) models suppose the series can be modelled as a number of distinct regimes, where the regimes are characterised by different conditional distributions of the process, each parameterised by an autoregression. For the self-exciting TAR (SETAR model), the regimes depend on observable lagged values of the process.

When there are two regimes, the process is in regime $i = 1$ at period t when $y_{t-d} \le r$, and otherwise $(y_{t-d} > r)$ in regime $i = 2$:

$$y_t = \phi_0^{\{i\}} + \phi_1^{\{i\}} y_{t-1} + \cdots + \phi_p^{\{i\}} y_{t-p} + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim \mathsf{IN}(0, \sigma_i^2), \quad i = 1, 2 \quad (1)$$

where the parameters super-scripted by $\{i\}$ may vary across regime. The orders of the autoregressions may differ across regimes (so that p is the maximum lag order and some of the $\phi_j^{\{i\}}$ may be zero for some i).

The Potter (1995) model, as reported in Hess and Iwata (1997), is:

$$y_t = -0.0071 + 0.302 y_{t-1} - 0.600 y_{t-2} + 0.028 y_{t-5} + \varepsilon_{1t} \quad \text{if } y_{t-2} \le 0$$

$$y_t = 0.0039 + 0.326 y_{t-1} + 0.195 y_{t-2} - 0.060 y_{t-5} + \varepsilon_{2t} \quad \text{if } y_{t-2} > 0$$

where the regime 1 standard error, $\sigma_1 = 0.0121$ and $\sigma_2 = 0.0088$, with $\mathsf{E}[\varepsilon_{1t}\varepsilon_{2t}] = 0$. The estimation period was 1947:1 to 1994:2, and y_t is the difference of the log of GDP. The data simulated from this model is multiplied by 100.

The Tiao and Tsay (1994) SETAR model has four regimes:

$$y_t = -0.015 - 1.076 y_{t-1} + \varepsilon_{1t} \quad \text{if } y_{t-1} \le y_{t-2} \le 0$$

$$y_t = -0.006 + 0.630 y_{t-1} - 0.756 y_{t-2} + \varepsilon_{2t} \quad \text{if } y_{t-1} > y_{t-2} \text{ and } y_{t-2} \le 0$$

$$y_t = 0.006 + 0.438 y_{t-1} + \varepsilon_{3t} \quad \text{if } y_{t-1} \le y_{t-2} \text{ and } y_{t-2} > 0$$

$$y_t = 0.004 + 0.443 y_{t-1} + \varepsilon_{4t} \quad \text{if } y_{t-1} > y_{t-2} > 0$$

and $\sigma_1 = 0.0062$, $\sigma_2 = 0.0132$, $\sigma_3 = 0.0094$ and $\sigma_4 = 0.0082$, and the errors are all independent and normally distributed. Regime 1 $(y_{t-1} \le y_{t-2} \le 0)$ is marked by negative growth two periods ago $(t-2)$, worsening in period $(t-1)$, and is characterised by an explosive root to bring the economy out of recession. Regime 2 implies negative growth in $t-2$, but improving in $t-1$. Regimes 3 and 4 are similar, and are operative when $t-2$ growth was positive and either slowed in $t-1$ or accelerated. The estimation period was 1947:1 to 1991:1, for the difference of the log of GNP. (The data simulated from this model is multiplied by 100).

Van Dijk and Franses (1999) present a four-regime smooth TAR (STAR – smooth transition autoregressive) model:

$$y_t = (0.394 + 0.460y_{t-1} + 0.092y_{t-2})(1 - F(\Delta y_{t-1}))(1 - F(CDR_{t-2}))$$
$$+ (-.121 + 0.442y_{t-1} + 0.3462y_{t-2})F(\Delta y_{t-1})(1 - F(CDR_{t-2}))$$
$$+ (0.360 - 0.530y_{t-1} - 0.963y_{t-2})(1 - F(\Delta y_{t-1}))F(CDR_{t-2})$$
$$+ (-0.019 + 0.744y_{t-1} - 0.235y_{t-2})F(\Delta y_{t-1})F(CDR_{t-2}) + \varepsilon_t$$

where: $F(\Delta y_{t-1}) = (1 + \exp[-500(\Delta y_{t-1} - 0.250)/\sigma_{\Delta y_{t-1}}])^{-1}$, $F(CDR_{t-2}) = (1 + \exp[-500(CDR_{t-2} - 0.064)/\sigma_{CDR_{t-2}}])^{-1}$, $\varepsilon_t \sim \mathsf{IN}(0, 0.867^2)$, the estimation period is 1947:1 to 1995:2, and y_t is the difference of the log of GDP multiplied by 100. With two transition functions (with transition variables Δy_{t-1} and CDR_{t-1}), the model is viewed as distinguishing between four regimes, generated by combinations of whether the level of output is above or below its previous high (the CDR variable), and whether growth is increasing or decreasing (the change in the growth rate, Δy_t). The CDR variable is the 'current depth of recession' variable of Beaudry and Koop (1993), except that $CDR_t = \max\{X_{t-j}\}_{j\geq 1} - X_t$, so the maximum is over past values only and not the current: see the discussion in section 2.3.

2.2. Markov-switching models

The Markov-switching autoregressive (MS-AR) model of Hamilton (1989) supposes contraction and expansionary regimes are different conditional distributions of the growth rate of real GNP (as for the SETAR model), but where the regime depends upon an unobserved state variable that follows a Markov chain. The Hamilton (1989) model of the US business cycle fits a fourth-order autoregression ($p = 4$) to the quarterly percentage change in US real GNP from 1953 to 1984:

$$y_t - \mu(s_t) = \alpha_1(y_{t-1} - \mu(s_{t-1})) + \cdots + \alpha_4(y_{t-4} - \mu(s_{t-4})) + \varepsilon_t, \tag{2}$$

where $\varepsilon_t \sim \mathsf{IN}(0, 0.768^2)$, $\{\alpha_1, \ldots, \alpha_4\} = \{0.014, -0.058, -0.247, -0.231\}$ and the conditional mean $\mu(s_t)$ switches between two states (although because of the autoregressive structure, the intercept switches between 32 values):

$$\mu(s_t) = \begin{cases} \mu_1 = -0.3577 & \text{if } s_t = 0 \text{ ('contraction')}, \\ \mu_2 = 1.522 & \text{if } s_t = 1 \text{ ('expansion')}. \end{cases}$$

with transition probabilities $p_{00} = 0.7550$, $p_{11} = 0.9049$ where:

$$p_{ij} = \Pr(s_{t+1} = j \mid s_t = i), \quad \sum_{j=1}^{2} p_{ij} = 1 \ \forall i, j \in \{0, 1\}. \tag{3}$$

Some authors have relaxed the assumption of fixed transition probabilities p_{ij}, and models with time-varying and duration-dependent transition probabilities have been considered (see, for example, Diebold, Rudebusch and Sichel, 1993, Diebold, Lee and Weinbach, 1994, Filardo, 1994, Lahiri and

Wang, 1994, and Durland and McCurdy, 1994). The former are modelled as logistic functions (to bound the probabilities between 0 and 1) of economic variables. When applied to modelling US GNP, the latter indicate that the probability of transition out of recession is increasing in the duration of the recession (see Durland and McCurdy, 1994, Filardo, 1994). The model we include in our study is that of Durland and McCurdy (1994), which is as the Hamilton (1989) model given above, but with $\varepsilon_t \sim \mathsf{IN}(0, 0.761^2)$, $\{\alpha_1, \ldots, \alpha_4\} = \{-0.017, -0.092, -0.255, -0.246\}$,

$$\mu(s_t) = \begin{cases} \mu_1 = -0.448 & \text{if } s_t = 0, \\ \mu_2 = 1.594 & \text{if } s_t = 1. \end{cases}$$

and:

$$\Pr[s_t = 0 \,|\, s_{t-1} = 0 \text{ AND } D_{t-1} = d] = \frac{\exp(6.516 - 1.348d)}{1 + \exp(6.516 - 1.348d)}$$

$$\Pr[s_t = 1 \,|\, s_{t-1} = 1 \text{ AND } D_{t-1} = d] = \frac{\exp(4.305 - 0.243d)}{1 + \exp(4.305 - 0.243d)}$$

Here, D_{t-1} is the number of periods the system has been in the current state (up to some maximum, set to 9). The coefficient of d in p_{11} is much smaller (robust standard error of 0.282, thus insignificant) than that on d in p_{00}, implying duration dependence in state 0 (such that $p_{00} = 0.0036$ for $d \geq 9$, compared to 0.994 for $d = 1$) but not state 1. The estimation period is identical to that of Hamilton (1989).

McCulloch and Tsay (1994) estimate a two-regime model in which the intercepts and autoregressive parameters are allowed to vary across regimes:

$$y_t = \begin{cases} -0.420 + 0.316y_{t-1} + 0.628y_{t-2} - 0.073y_{t-3} - 0.097y_{t-4} + \varepsilon_{0t} \\ \qquad \text{if } s_t = 0 \text{ (`contraction')}, \\ 0.909 + 0.265y_{t-1} + 0.029y_{t-2} - 0.126y_{t-3} - 0.110y_{t-4} + \varepsilon_{1t} \\ \qquad \text{if } s_t = 1 \text{ (`expansion')}. \end{cases}$$

where $p_{00} = 0.714$, $p_{11} = 0.882$, and $\sigma_0 = 1.017$, $\sigma_1 = 0.816$. The model is estimated for 1947:2 to 1991:1, and by the Gibbs sampler because the authors wish to conduct a Bayesian analysis of the model. The coefficients given above are the posterior means, taken from McCulloch and Tsay (1994, Table 1, p534, Model (7)). This model is a univariate special case of the switching regression model of Lindgren (1978): see Tyssedal and Tjøstheim (1988) for an application.

Finally, we include the three-regime MS model of Clements and Krolzig (1998) (Boldin (1996) also estimates a three-regime model). This has a shifting intercept term and a heteroskedastic error term (and is denoted by the label MS3).

$$y_t = \mu(s_t) + \sum_{k=1}^{4} \alpha_k y_{t-k} + \varepsilon_t, \tag{4}$$

Table 1. Model labels and sources

Label	Source	Description
SET2	Potter (1995)	2-regime SETAR model
SET4	Tiao and Tsay (1994)	4-regime SETAR model
STAR	Van Dijk and Franses (1999)	4-regime STAR with output growth and CDR transition variables
MS2	Hamilton (1989)	2-regime MS-in-mean, homoscedastic model
MS2dd	Durland and McCurdy (1994)	2-regime MS-in-mean with duration dependence
MSA2	McCulloch and Tsay (1994)	2-regime MS model, with slopes and means changing
MS3	Clements and Krolzig (1998)	3-regime MS-in-intercept, heteroscedastic model
CDR	Beaudry and Koop (1993)	Current depth of recession model
F&C	Pesaran and Potter (1997)	Floor and ceiling model

where $\varepsilon_t \sim \mathsf{IN}(\sigma^2(s_t))$ and $s_t \in \{1,2,3\}$ is generated by a Markov chain. The intercepts are $\{\mu_1, \mu_2, \mu_3\} = \{-0.063, 0.866, 1.444\}$, the slopes $\{a_1, \ldots, a_4\} = \{0.013, -0.023, -0.128, -0.056\}$, and $\{\sigma_1^2, \sigma_2^2, \sigma_3^2\} = \{0.772, 0.118, 0.405\}$. The transition matrix is:

$$P = \begin{bmatrix} 0.848 & 0.022 & 0.130 \\ 0.075 & 0.925 & 0 \\ 0 & 0.090 & 0.910 \end{bmatrix}.$$

The three-state model captures Sichel's depiction of post-War US business cycles as consisting of three phases: contraction, followed by high-growth recovery, and then a period of moderate growth: see Sichel (1994). This is consistent with $p_{12} \simeq 0$ (and $p_{23} = 0$), so that the economy moves directly from recession to high growth, and from high growth to moderate growth ($p_{31} = 0$). The estimation period is 1959:2 to 1996:2.

2.3. Endogenous threshold models

The 'current depth of recession' (CDR) model of Beaudry and Koop (1993), as given by Hess and Iwata (1997), is of the form:

$$y_t = 0.0016 + 0.447 y_{t-1} + 0.199 y_{t-2} + 0.351 CDR_{t-1} + \varepsilon_t$$

where the CDR variable is defined as:

$$CDR_t = \max\{X_{t-j}\}_{j \geq 0} - X_t,$$

X_t is the log of (the level of) output, $\varepsilon_t \sim N(0, 0.0093^2)$. CDR is either zero, when the current level of output is a historical maximum, or is the (positive) gap between the current level of output and the historical maximum. So, the deeper the current 'recession', the greater the impetus to growth. This model is sometimes interpreted as saying that positive shocks are more persistent than negative ones. Pesaran and Potter (1997) show that this model can be interpreted as a TAR model with a large number of regimes: the CDR variable offers a parsimonious way of allowing for such a model. Jansen and Oh (1999) compare the CDR model to a two-regime STAR. The estimation period for the model reported above is 1949:1 to 1992:4. Because $y_t = \Delta X_t$, the data simulated from this model is multiplied by 100.

The 'floor and ceiling' (F&C) model of Pesaran and Potter (1997) can be given a similar interpretation, and consists of:

$$\Delta y_t = -0.462\Delta y_{t-1} - 0.862CDR_{t-1} - 0.161OH_{t-1} + h_t\varepsilon_t \tag{5}$$

and the recursions:

$$F_t = \begin{cases} 1(y_t < r_f) & \text{if } F_{t-1} = 0 \\ 1(CDR_{t-1} + y_t < 0) & \text{if } F_{t-1} = 1 \end{cases}$$

$$CDR_t = \begin{cases} (y_t - r_f)F_t & \text{if } F_{t-1} = 0 \\ (CDR_{t-1} + y_t)F_t & \text{if } F_{t-1} = 1 \end{cases}$$

$$C_t = 1(F_t = 0)1(y_t > r_c)1(y_{t-1} > r_c)$$

$$OH_t = C_t(OH_{t-1} + y_t - r_c)$$

$$COR_t = 1(F_t + C_t = 0).$$

ε_t is standard normal, $h_t = (0.918COR_{t-1} + 1.173F_{t-1} + 0.685C_{t-1})$. $1(A)$ is an indicator function taking the value unity when the statement in brackets is true and zero otherwise, $\{r_c, r_f\}$ are the ceiling and floor threshold values, with values $\{-0.876, 0.539\}$. The two non-linear terms in (5), CDR_t and OH_t, are the 'current depth of recession' and an 'overheating variable'. Intuitively, the measure of the depth of the recession has the effect that as output falls below the floor threshold, pressure starts to build up (CDR_t becomes increasingly negative) for output to return to its previous levels. During these times $F_t = 1$, denoting that the floor regime is in operation. In addition, OH_t is an estimate of cumulated output growth above some ceiling, and is non-zero when $C_t = 1$. The corridor regime is in operation ($COR_t = 1$) when $F_t = C_t = 0$. Thus this construction allows different dynamics depending on which regime the economy is in.

The model is estimated over the period 1954:1 to 1992:4, on y_t, given by 100 times the difference of the log of GNP.

3. Conditional mean functions

Pagan and Ullah (1999, ch. 3) provide an excellent exposition of conditional moment estimation. We use a Nearest Neighbour (NN) estimator with uniform weights (see, e.g., Pagan and Ullah (1999, p. 89)), so that the conditional mean for $p = 1$ is:

$$\mathsf{E}[Y_t \mid Y_{t-1} = y] = \widetilde{m_1}(y) = k^{-1} \sum_{i=1}^{n} I_{ki}(y) y_i \tag{6}$$

where $I_{ki}(y) = 1$ if y_{i-1} is one of the k-nearest neighbours to y, and zero otherwise. n is the number of observations. We evaluate (6) at a grid of values y, as described below. When $p = 2$, this method entails calculating:

$$\mathsf{E}[Y_t \mid Y_{t-1} = y_1, Y_{t-2} = y_2] = \widetilde{m_2}(y_1, y_2) = k^{-1} \sum_{i=1}^{n} I_{ki}(y_1, y_2) y_i \tag{7}$$

where now $I_{ki}(y_1, y_2) = 1$ if $\{y_{i-1}, y_{i-2}\}$ is the k-th NN to $\{y_1, y_2\}$, using the Euclidean distance metric, and zero otherwise. (7) is evaluated at a two-dimensional grid of values, and the results plotted. In fact, for $p = 2$ instead of using the NN method we used LOESS local linear regression, since this may give more reliable estimates at points near the boundary of the parameter space. For $p = 1$, there was little difference between the NN and LOESS results, and we report the former. LOESS local linear regression is equivalent to weighted least squares estimation with the weights depending on the kernel function and on a nearest neighbour bandwidth (see, e.g., Cleveland, Grosse and Shyu (1993), Simonoff (1996)). We make use of the LOESS routine in S-PLUS 2000 (MathSoft, Inc). The algorithm is robust to outliers – using an iterative routine, less weight is given to observations with relatively large residuals. A tricube kernel is used to define the weights, and the neighbourhood size is set to 30%.

(6) estimates, for a given value of $Y_{t-1} = y$, the integral $\int y_t f_{Y_t|Y_{t-1}}(u|y)\,du$, where $f_{Y_t|Y_{t-1}}$ is of course $\int f_{Y_t|Y_{t-1},Y_{t-2}}(u|y_1, y_2)\,dy_2$, and $f_{Y_t|Y_{t-1},Y_{t-2}}$ is the density underlying (7). Conditional moment estimates of the form (6) are computationally less demanding and easier to compare across models, but whether interesting features are lost in the marginalization will depend on the nature of the data generating process. It seems sensible to regard this as an open question, rather than to only consider $p = 1$.

The data consist of realizations of $n = 25000$ observations generated from the nine non-linear and one linear models described in section 2. Data are generated for an additional 'burn-in' period, which is then discarded, as well as drawing the initial states of the Markov models from their unconditional distributions. The models' parameter values are taken to be the estimates provided in the literature, so these are based on different vintages of data and time periods, albeit that they are all on post War US GNP/GDP data. There is a case for estimating all the models on the same data, though we have chosen not to do so. This is because non-linear models are often sensitive to the precise estimation period,[3] and it would be unclear how to handle models whose parameter estimates altered to a 'large' extent relative to their original values. Taking the models as they are is reasonable if our interest is in the properties of the models themselves. However, a possibly more important point is that by effectively assuming that the models' parameter estimates are the population

[3] As an example, Boldin (1996) and Clements and Krolzig (1998, Section 4.3) find the original Hamilton (1989) two-regime MS-AR model is not robust to changes in the sample period, and propose three-regime models instead.

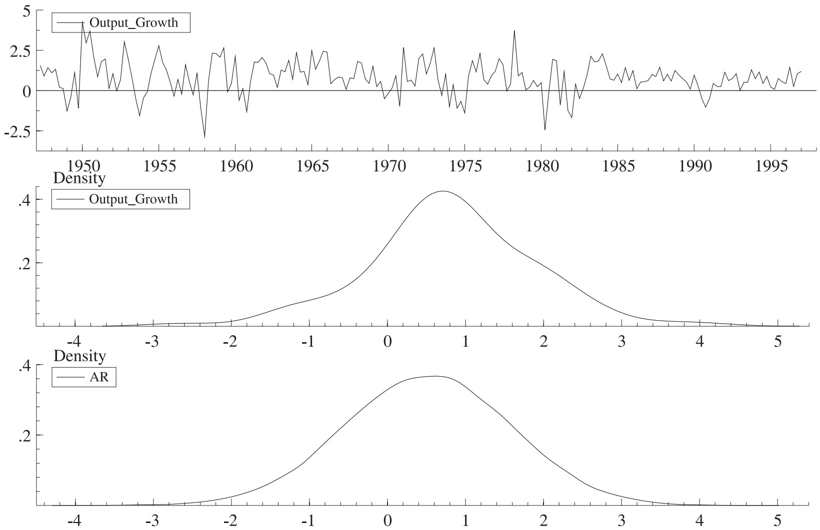

Fig. 1. US output growth 1947:2 to 1997:1 and data simulated from an AR(1) model of output growth.

values of the parameters we are neglecting parameter estimation uncertainty. Similarly, we are ignoring model specification uncertainty. These are important issues, but are beyond the scope of the study.

Actual data is used in our study in two ways. Firstly, the linear AR is estimated on data for the period 1947.1 to 1997.1. Since it is a low order autoregression (an AR(1) in the first difference) we would expect this to be quite robust to changes in the estimation period, so the issue of the precise estimation period does not arise. Secondly, the sample growth rates over this period are used to determine the grid of values at which the conditional mean functions will be estimated. Here again, the range of growth rates is unlikely to be excessively sensitive to the precise sample period.[4]

Figure 1 plots the actual growth rates, the estimated density for the data, and the estimated density of the data simulated from the AR model.[5] For the last two, the scale of the horizontal axis is the same to aid comparison. It is clear that the symmetric AR model density is at odds with the skewness in the actual data. Non-parametric tests for business cycle asymmetries based on the coefficient of skewness for the detrended series (and the difference of the detrended series) have been applied by a number of authors, see, e.g., Sichel (1993, p. 227–8), and Clements and Krolzig (2000) develop parametric tests for asymmetries based on the MS-AR model. Figure 2 and table 2 summarise some properties of the simulated and actual quarterly changes series (more precisely, one hundred times the first difference of the natural logarithm of constant price GNP). Table 1 collects together the model labels and descriptions for ease of reference.

[4] What we would expect to turn on the precise sample period would be, say, transition probabilities between regimes, which will depend on whether particular recessions are included or not.

[5] The data were plotted in GiveWin using the default options for the densities: see Doornik and Hendry (1996).

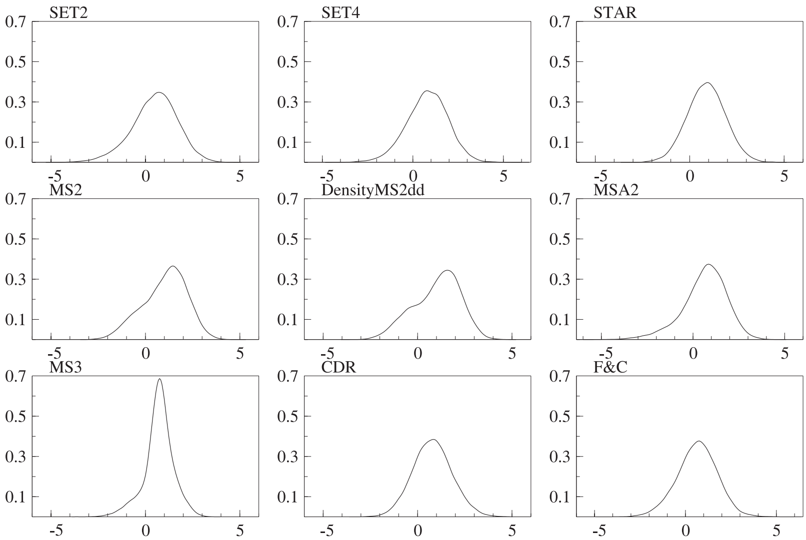

Fig. 2. Densities of data simulated from non-linear models of output growth.

Table 2. Actual and simulated data characteristics

	SET2	SET4	STAR	MS2	MS2dd	MSA2	MS3	CDR	F&C	AR	Data
mean	0.50	0.79	0.88	1.0	1.0	0.63	0.70	0.79	0.64	0.51	0.81
STD	1.2	1.1	1.0	1.2	1.2	1.2	0.80	1.0	1.1	1.1	1.1
min	−4.9	−4.4	−3.3	−3.1	−3.7	−6.0	−3.7	−3.3	−4.6	−4.0	−2.9
max	4.7	6.0	4.9	5.0	4.9	4.7	3.7	4.8	6.1	4.7	4.2
$\% < -2$	3.0	1.2	0.24	0.65	0.94	3.2	0.38	0.30	1.1	0.98	
$\% > 3$	1.2	2.1	1.7	2.5	3.1	1.1	0.11	2.0	1.4	0.91	

Notes: The rows are the means, standard deviations, and minimum and maximum observations. The row $\% < -2$ denotes the % of observations less than -2, and $\% > 3$ the % of observations greater than 3. The interval $(-2, 3)$ contains approximately 95% of the actual data.

From the figure, it is apparent that some of the marginal densities are left skewed, and there is some variation in the left tails.

We chose to evaluate (6) at a grid of 100 equidistant values ranging from -2 to 3: this interval contains approximately 95% of the actual values of the quarterly growth rate change over the period 1947.1 to 1997.1. We set k to 5% of the observations, i.e., 1250. Too large a choice will 'over-smooth', too small will fail to bring out salient features. From table 2 it is apparent that the estimates at the lower end point may be unreliable for some of the models (e.g., the STAR, MS3 and CDR) where less than 1% of the simulated observations are smaller. Where there are few observations, the conditional means will be flat at these points (because such points will tend to share the same nearest neighbours), and this has to be borne in mind when interpreting the conditional mean estimates.

(7) is evaluated at a two dimensional grid of values with the same range. The problem of few simulated observations near some points at which we wish

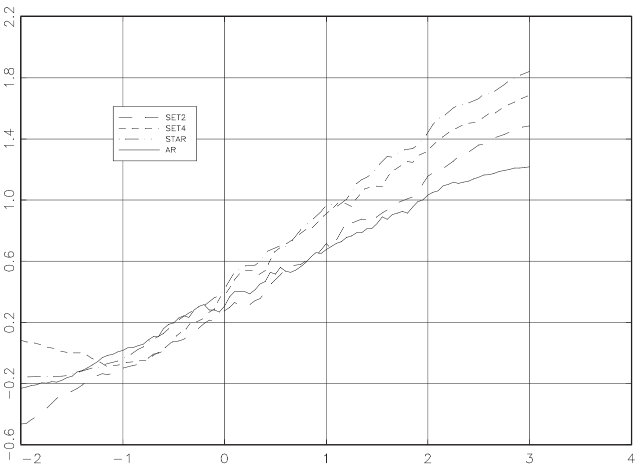

Fig. 3. Threshold models' conditional means. The ordinate is the estimate of $E[Y_t \mid Y_{t-1} = y]$ and the abscissa is Y_{t-1}.

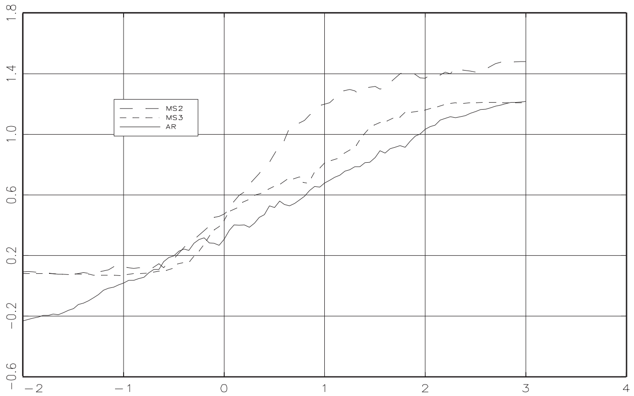

Fig. 4. Markov models' conditional means. The ordinate is the estimate of $E[Y_t \mid Y_{t-1} = y]$ and the abscissa is Y_{t-1}.

to evaluate the conditional expectation is exacerbated relative to $p = 1$. For example, two consecutive quarters of declines in output of 2% points are historically infrequent in the post War US economy, and are unlikely to be generated very often by our models. We use LOESS, as explained above.

4. Results

The estimated conditional mean functions (6) are displayed in four figures, figs. 3 to 6, where the grouping is chosen to highlight the impact on the mean functions of particular features of the models. Consider firstly the threshold models in fig. 3 – the two and four regime models (SET2 and SET4) and the 4-regime

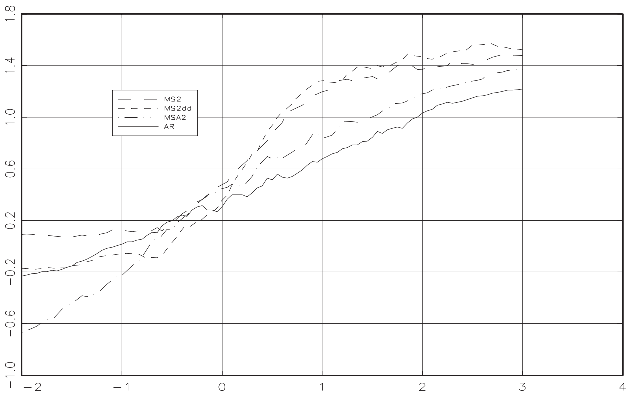

Fig. 5. Markov models' conditional means. The ordinate is the estimate of $E[Y_t \mid Y_{t-1} = y]$ and the abscissa is Y_{t-1}.

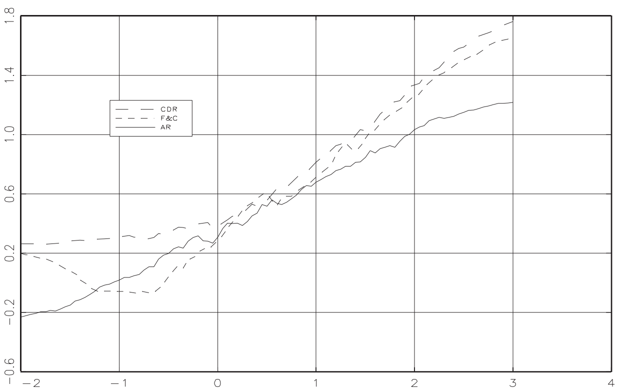

Fig. 6. Endogenous threshold models' conditional means. The ordinate is the estimate of $E[Y_t \mid Y_{t-1} = y]$ and the abscissa is Y_{t-1}.

STAR model. The AR is included for comparison. The SET2 conditional mean is more negative than that of the AR model for negative $t-1$ output growth, while the STAR model is similar to the AR for $y_{t-1} < 0$ but considerably steeper for $y_{t-1} > 0$. The STAR model would predict growth of nearly $1\frac{1}{2}$ in response to growth in the previous period of 2%, approximately $\frac{1}{2}$% more than the AR model, and 0.3% more than SET2. SET4 also suggests strong growth, and a positive response to increasingly negative rates.

Many business cycle analysts use simple sequencing rules, such as the occurrence of two periods of negative growth, followed by two periods of positive growth, to date a trough, because rules of this sort approximately correspond to the Bry and Boschan (1971) algorithm for dating cycles (without any censoring to rule out short cycles or phases), and thus to the NBER dating chronology more generally – see, e.g., Harding and Pagan (2001). Thus, the SET4 model is *less* likely to predict recessions than the other models. Despite the graph never going negative, recessions are not ruled out. The graph depicts the conditional mean while the ocurrence of a recession depends on the actual observations (which include the disturbances).

Figure 4 depicts the MS2 and MS3 models, with the AR as a reference. The MS models are similar for $y_{t-1} < 0$, and flat at 0.1 for $y_{t-1} < -\frac{1}{2}$, so that neither model indicates that two consecutive declines in output are likely. But the MS models differ considerably for $y_{t-1} > 0$: the MS2 line rises much more sharply, and for $y_{t-1} = 1$ the MS2 model predicts a response $\frac{1}{2}$% higher. Figure 5 considers two other extensions to the basic MS model. These are duration dependence of recessions, MS2dd, and allowing the AR parameters to depend on the regime, MSA2. The MS2 and AR models are plotted as well for convenience. Firstly, notice that the MSA2 model is similar to the SET2 model in fig. 3. This is perhaps unsurprising, because both are two regime models that allow all the parameters to change, and differ only as to the way in which regime switches are generated. So relative to the MS2 model, allowing the AR coefficients to change (the MSA2), increases the likelihood of two consecutive falls in output. Allowing for duration dependence in recessions has little impact for $y_{t-1} > 0$ (the MS2 and MS2dd lines are close together), but increases the likelihood of two consecutive falls, because the conditional mean is negative for $y_{t-1} = \{-2, -\frac{1}{2}\}$. This is at first sight paradoxical, but note that with duration dependence, the probability of moving out of recession the period after entering it is less than when duration dependence is absent.

Finally, consider fig. 6 that depicts the CDR and F&C models. The models are similar for $y_{t-1} > 0$ and indicate a response a little higher than for SET2. But the CDR model appears to pivot at zero and declines only a little thereafter, to predict growth of a $\frac{1}{4}$% at $y_{t-1} = -2$. By contrast, the F&C model continues to decline until $y_{t-1} = -\frac{3}{4}$, where it is just negative, and then turns up.

How much information is lost in considering (6) rather than (7)? The conditional mean functions for $p = 2$ estimated by LOESS are given in figs. 7 and 8. For the linear AR model the surface does not vary in the Y_{t-2} direction, because the correlation between Y_t and Y_{t-1} does not depend upon the value of Y_{t-2}. More unexpectedly, the same appears to be true of the STAR model: we suspect the smooth transition between regimes, that weights the regimes together, accounts for this. For the SET2 model, however, there is a trough in the surface at around $Y_{t-2} = 0$, so that for values of Y_{t-1} between $+1$ and -1, say, the expected value of output growth in period t is increasing as Y_{t-2} moves below, or above, zero. The importance of the $t-2$ level of output growth is due to the two-period delay in determining the regime in this model, and the large negative autoregressive term in the 'recession' regime at lag 2, which explains why the more negative the output growth in $t-2$, the greater the fillip to period t. The SETAR model with a finer division of regimes, SET4, indicates that the increase in the conditional mean as Y_{t-1} increases from, say -1, is much faster than for SET2, and is less dependent on Y_{t-2}.

Consider now the four MS models (MS2, MS2dd, MSA2, and MS3). There are a number of interesting points to note. Firstly, we remarked upon the similarity of the $p = 1$ conditional mean plots for SET2 and MSA2. Comparing the figures for $p = 2$, it is evident that this is due to averaging over Y_{t-2}: the plots for $p = 2$ are quite different. The surface for the MSA2 model increases approximately linearly in both directions, though at a faster rate in Y_{t-1}. Secondly, allowing for changing AR parameters appears to remove the concavity in the MS2 and MS2dd models, where the conditional mean is relatively flat

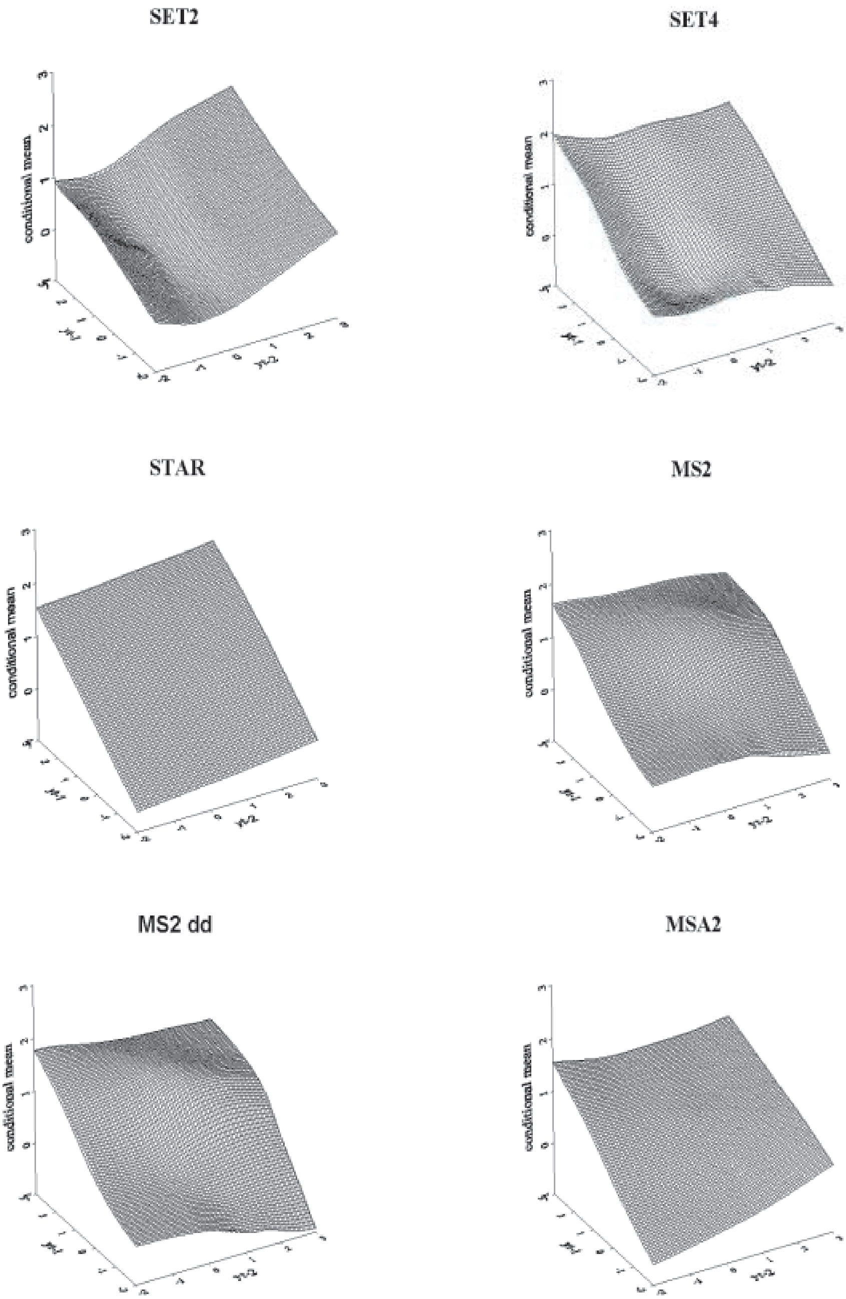

Fig. 7. Conditional mean functions of $p = 2$ estimated by LOESS

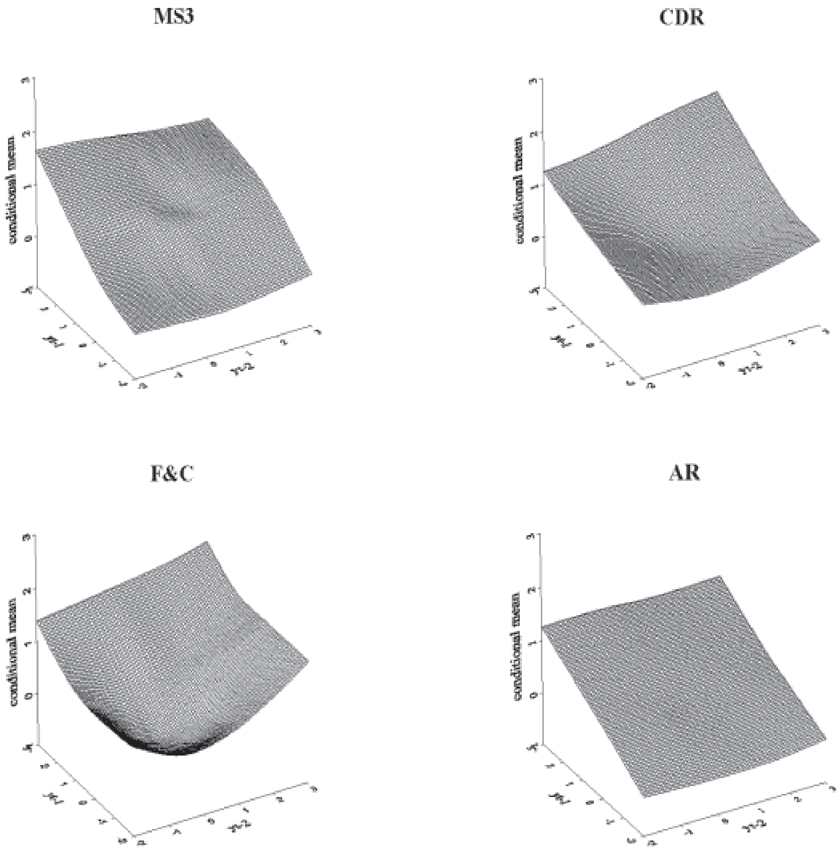

Fig. 8. Conditional mean functions for $p = 2$ estimated by LOESS

for high and low pairs of $\{Y_{t-1}, Y_{t-2}\}$, compared to a steep gradient around $\{0, 0\}$. The effect of allowing for duration dependence is not easy to discern (MS2dd), while the third regime (MS3) causes a slight hollow at high values of $\{Y_{t-1}, Y_{t-2}\}$.

The CDR model suggests that output growth rises more quickly as Y_{t-1} increases at high levels of Y_{t-2}. The F&C model surface is broadly similar, although there is some evidence of a valley running in the direction of growth of around 2% in Y_{t-1} and high growth in Y_{t-2}, down to low (or negative) growth in Y_{t-1} and Y_{t-2}. In general, the models with three or more regimes (such as the SET4, MS3, CDR and F&C) have richer dynamics, as one would expect.

5. Conclusions

We have shown that the rival non-linear models of US output growth can imply quite different behaviour for the conditional mean given the previous growth rate, and that this would not necessarily be apparent from a cursory examination of the models themselves. In particular, the models would appear to differ in terms of the likelihood of two consecutive declines of output occurring, because in some cases the conditional means, evaluated at all data-supported values of the previous growth rate are everywhere positive, while for other other models negative conditional means arise. Still others suggest a stronger fillip to output growth the greater the fall last period. The projection on to the previous two periods' output growth rates suggests that in some cases the expectation conditional on only the last period may obscure interesting differences in the models' predictions of output growth rates.

A natural question which we have so far avoided is which of the models is the best? As we note in the introduction, this partly depends on how 'best' is defined. A sceptical view might be that on fifty years of quarterly data, of which only around 20% of the observations record declines in output (where we see some of the largest differences in model behaviour), and given that the evidence for non-linearity is relatively weak (e.g., Hansen, 1992, 1996a), it would be very optimistic to expect formal testing to be able to discriminate between the subtle differences in specification, that nevertheless contribute to appreciable differences in the conditional mean estimates across models.

References

Beaudry P, Koop G (1993) Do recessions permanently affect output: Journal of Monetary Economics 31:149–163

Boldin MD (1996) A check on the robustness of Hamilton's Markov switching model approach to the economic analysis of the business cycle. Studies in Nonlinear Dynamics and Econometrics 1:35–46

Bry G, Boschan C (1971) Cyclical analysis of time series: Selected procedures and computer programs. New York: NBER

Clements MP, Krolzig H-M (1998) A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP. Econometrics Journal 1:C47–75

Clements MP, Krolzig H-M (1999) Modelling business cycle features using switching regime models. mimeo, Department of Economics, University of Warwick

Clements MP, Krolzig H-M (2000) Business cycle asymmetries: Characterisation and testing based on Markov-switching autoregressions. mimeo, Department of Economics, University of Warwick

Clements MP, Smith J (1999) A Monte Carlo study of the forecasting performance of empirical SETAR models. Journal of Applied Econometrics 14:124–141

Cleveland WS, Grosse E, Shyu WM (1993) Local regression models. In Chamber JM, Hastie TJ (eds.), Statistical Models in S, pp. 309–376: London: Chapman & Hall

Davies RB (1977) Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 64:247–254

Davies RB (1987) Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 74:33–43

Diebold FX, Lee JH, Weinbach GC (1994) Regime switching with time-varying transition probabilities. In Hargreaves C (ed.), Non-stationary Time-series Analyses and Cointegration, pp. 283–302: Oxford: Oxford University Press

Diebold FX, Rudebusch GD, Sichel DE (1993) Further evidence on business cycle duration dependence. In Stock J, Watson M (eds.), Business Cycles, Indicators, and Forecasting, pp. 255–280: Chicago: University of Chicago Press and NBER

Doornik JA, Hendry DF (1996) GiveWin: An interactive empirical modelling program. London: Timberlake Consultants Press

Durland JM, McCurdy TH (1994) Duration dependent transitions in a Markov model of U.S. GNP growth. Journal of Business and Economic Statistics 12:279–288

Filardo AJ (1994) Business-cycle phases and their transitional dynamics. Journal of Business and Economic Statistics 12:299–308

Gallant AR, Rossi PE, Tauchen G (1993) Nonlinear dynamic structures. Econometrica 61:871–907
Galvão AB (2000) Univariate non-linear time series models and business cycle stylised facts. Department of Economics, mimeo, Warwick University
Hamilton JD (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57:357–384
Hansen BE (1992) The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP. Journal of Applied Econometrics 7:S61–S82
Hansen BE (1996a) Erratum: The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP. Journal of Applied Econometrics 11:195–198
Hansen BE (1996b) Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64:413–430
Harding D, Pagan AR (2001) Dissecting the cycle: A methodological investigation. Journal of Monetary Economics. Forthcoming
Hess GD, Iwata S (1997) Measuring and comparing business-cycle features. Journal of Business and Economic Statistics 15:432–444
Jansen DW, Oh W (1999) Modeling nonlinearity of business cycles: Choosing between the CDR and STAR models. Review of Economics and Statistics 81:344–349
King RG, Plosser CI (1994) Real business cycles and the test of the Adelmans. Journal of Monetary Economics 33:405–438
Koop G, Pesaran MH, Potter SM (1996) Impulse response analysis in nonlinear multivariate models. Journal of Econometrics 74:119–147
Lahiri K, Wang JG (1994) Predicting cyclical turning points with leading index in a Markov switching model. Journal of Forecasting 13:245–263
Lindgren G (1978) Markov regime models for mixed distributions and switching regressions. Scandinavian Journal of Statistics 5:81–91
Luukkonen R, Saikkonen P, Teräsvirta T (1988) Testing linearity against smooth transition autoregressive models. Biometrika 75:491–499
McCulloch RE, Tsay RS (1994) Bayesian analysis of autoregressive time series via the Gibbs sampler. Journal of Time Series Analysis 15:235–250
Pagan A, Ullah A (1999) Nonparametric econometrics. Cambridge: Cambridge University Press
Pagan AR (1997a) Policy, theory and the cycle. Oxford Review of Economic Policy 13:19–33
Pagan AR (1997b) Towards an understanding of some business cycle characteristics. Australian Economic Review 30:1–15
Pagan AR (1999a) Bulls and bears: A tale of two states. mimeo, Australian National University
Pagan AR (1999b) The getting of macroeconomic wisdom. mimeo, Australian National University
Pesaran MH, Potter SM (1997) A floor and ceiling model of US output. Journal of Economic Dynamics and Control 21:661–695
Potter S (1995) A nonlinear approach to US GNP. Journal of Applied Econometrics 10:109–125
Sichel DE (1993) Business cycle asymmetry. Economic Inquiry 31:224–236
Sichel DE (1994) Inventories and the three phases of the business cycle. Journal of Business and Economic Statistics 12:269–277
Simonoff JS (1996) Smoothing methods in statistics. New York: Springer-Verlag
Tiao GC, Tsay RS (1994) Some advances in non-linear and adaptive modelling in time-series. Journal of Forecasting 13:109–131
Tyssedal JS, Tjøstheim D (1988) An autoregressive model with suddenly changing parameters and an application to stock market prices. Applied Statistics 37:353–369
Van Dijk D, Franses PH (1999) Modelling multiple regimes in the business cycle. Macroeconomic Dynamics 3:311–340
van Dijk D, Franses PH, Boswijk HP (2000) Asymmetric and common absorption of shocks in nonlinear autoregressive models. Mimeo, Econometric Institute Research Report 2000-01/A, Erasmas University, Rotterdam
Watson WM (1994) Business cycle durations and postwar stabilization of the U.S. economy. American Economic Review 84:24–46