ISSN 1440-771X
0 7326 1096 6

MONASH UNIVERSITY
AUSTRALIA

DEPARTMENT OF ECONOMETRICS
AND BUSINESS STATISTICS

Non-linear Modelling of the Australian Business Cycle Using a Leading Indicator

Roland G Shami and Catherine S Forbes

Working Paper 5/2002

Non-linear Modelling of the Australian Business Cycle using a Leading Indicator

Roland G. Shami and Catherine S. Forbes

Abstract:

This paper develops a new non-linear model to analyse the business cycle by exploiting the relationship between the asymmetrical behaviour of the cycle and leading indicators. The model proposed is an innovations form of the structural model underlying simple exponential smoothing that is augmented by a latent Markov switching process. Furthermore, the probabilities that drive the Markov process vary with the growth of the leading indicator. The proposed model is used to analyse the Australian business cycle using the gross domestic product as a proxy and the industrial materials prices index as the exogenous leading indicator influencing the transition probabilities. Model parameters are estimated using a Gibbs sampling algorithm and subsequently used for forecasting purposes.

Key words: Structural model - Markov switching regime - Gibbs sampling - Business cycle - Leading indicator.

JEL Classification: C11 – C15 – C22 – C51 - E32.

1. Introduction

Business cycles are loosely defined as recurrent sequences of alternating expansion and contraction in economic activity, which is identified by the movement of different variables such as employment, retail sales, industrial production and others. Some variables are much more sensitive to the movement of business cycles than are others. Some tend to move early (leading indicators) and some tend to move late (lagging indicators) while others are roughly coincident indicators. An example of the coincident indicators is the gross national product (GNP) and the gross domestic product (GDP), which are often used as proxies of the business cycle to analyse the situation of the economic activity (Hamilton, 1989; Luginbuhl and De Vos, 1999). Simple deterministic models, such as the popular Bry-Boschan (1971) procedure that defines a recession to occur when there are two consecutive declines in GDP, lacks predictive ability because of the delayed announcement of the recession. This delay is due to the gap in time between the supposedly beginning of the recession and the publication of the data.

The same delay happens with the National Bureau of Economics Research (NBER) announcement in the United States. Recently, in November 2001 the NBER announced that a recession had begun in March 2001. This late announcement, eight months after the selected date, is not uncommon because the committee responsible for the dating of business cycles need to validate the actions of some indicators by the actions of other indicators. In this paper, we are assessing the potential usefulness of the leading indicators in forecasting the business cycle.

This paper develops a new non-linear model to analyse the business cycle. During the last two decades, there has been conformity among most of the researchers that non-linear time series modelling would improve forecasts and produce a richer notion of business cycle dynamics than linear time series models

(Teräsvirta and Anderson, 1992; Beaudry and Koop, 1993; Potter, 1995; …). This is due to the inability of the linear models to capture the expansionary and recessionary phases in the business cycle that display asymmetric behaviour in the sense that the arrival of a recession is prompt, while the recovery from a recession is prolonged. Additionally, often two cycles appear to have neither the same amplitude nor the same duration. As such, non-linear models are used in preference to linear models to characterise this distinction. In this paper we propose a new family of models based on the innovations form of the structural time series model that is augmented by a latent binary switching variable.

The characterisation of an economic time series using linear structural models is based on a traditional decomposition of the observed series into level, growth, seasonal and random components (Harvey, 1984). These unobserved components are assumed to evolve dynamically according to a linear relationship, traditionally made stochastic by the inclusion of an additive error term that is uncorrelated with the observation error. Statistical analysis of linear structural models requires writing them in state space form and using the Kalman filter to estimate them. Another equally general state space framework involves only a single source of error (Snyder, 1985; Ord, Koehler and Snyder, 1997). Called the innovations form by Aoki (1987), the calculation of the likelihood function for this model is made easier by using exponential smoothing methods rather than the Kalman filter. It also has a more direct equivalence relationship to the popular autoregressive integrated moving average (ARIMA) models than does the traditional linear structural model (Shami and Snyder, 1998). Linear state space models in both the traditional form (Harvey, 1985; Watson, 1986; and Clark, 1987) and the innovations form (Aoki, 1988 and 1993) have been used to characterise economic time series. Notably, Harvey (1985) used the linear structural model on US GNP data to analyse the business cycle.

The switching initiative was introduced in economic analysis by Hamilton (1989) to define changes between fast and slow growth regimes in the economy. He proposed a new univariate model and applied it to characterise the US business cycle using US GNP data. His model uses a simple autoregressive structure to characterise the evolution of the observed series whose conditional mean is determined by a latent, binary Markov switching variable that takes a value of unity during expansionary periods and a value of zero during recessions. Hamilton also provided an algorithm for estimating the probability of a recession at each time period based on a maximum likelihood approach. Since that time, several other authors have investigated modifications to the model specification (Lam, 1990; Hansen, 1992; Kim, 1994), computation of the recession probabilities (Albert and Chib, 1993) and the application of the models to various other data sources (Cecchetti et al, 1990; Hamilton and Lin, 1996).

However, the stability of Hamilton model using the likelihood approach is a matter of concern. Boldin (1996) observed a breakdown of the Hamilton model for data, which includes the end of World War II and the Korean War. While, Kim and Nelson (1999) found that Hamilton's model fails to provide reasonable inferences on the probabilities of a recession or a boom when Hamilton's original data set is extended until 1992. To correct for this, they added a dummy variable from 1983 to account for a structural break in the growth rate. This idea is in line with our proposed model, which allows for "structural breaks" at each time through a change in the level. Others like Krolzig (2001) and Luginbuhl and De Vos (1999) found structural breaks in the growth rate.

The proposed model, which we will call the switching structural model (SSM), is described as a non-linear structural model that includes an unobserved level component and an unobserved switching drift component. The level is estimated by simple exponential smoothing where at a certain time *t*, it is expressed as a weighted sum of the observed data at time *t* and the level at time *t*-1. The drift is represented by a variable that switches between two values that represent the expected rates of growth during an expansion and a recession. These values evolve according to a Markov chain of order one with constant transition probabilities.

The SSM proves to be satisfactory in estimating the business cycle by modelling a coincident indicator; still, the predicted values depend on the regime state at the end of the observation period and the estimated common values of the transition probabilities. One of the primary aims of this paper is to develop an improved forecasting model by exploiting the characteristic of a leading indicator. Thus, we extend the SSM by relaxing the assumption of constant transition probabilities. We introduce the switching structural model with varying transition probabilities (SSMVP) that relates to a leading economic indicator. One of the potential uses of these indicators is that the persistence of an expansion or a recession depends on both the state of the system at one period earlier and the leading indicator which has the ability to provide a direct measure of the expectations of the economic agents. Also, it helps in improving the forecast and predictive accuracy around the turning points. What is probably of greater interest is as these variables vary across time, the persistence of business cycle phases will vary and thus affect the expected duration of how long an expansion or a recession will last.

Hamilton (1990) noted that one of the reasons for the instability of his model might be owed to the computational difficulty of maximising numerically an often ill behaved likelihood surface with respect to a large number of unknown parameters. As such, we provide a new approach to conduct a Bayesian analysis of the proposed models. The Gibbs sampling (Gelfand and Smith, 1990) based algorithm builds on the work of Forbes, Snyder and Shami (2000), who demonstrate the use of Monte Carlo composition to compute Bayesian posterior parameter and forecasting distributions for the linear structural model based on the innovations form. Others, notably Albert and Chib (1993), Kim and Nelson (1999) and Luginbuhl and De Vos (1999) have used Bayesian methods on various traditional switching models.

The plan of the paper is as follows. We detail the proposed non-linear model in section 2 and demonstrate properties of the model. A Bayesian estimation algorithm based on a Gibbs sampler scheme is presented in section 3. In section 4, the new proposed model will be used to analyse the Australian business cycle. As there is no authority in Australia similar to the NBER in the USA to use as a benchmark or a point of reference, the recession and expansion dates are calculated using the Bry-Boschan procedure, which will be used for the purpose of comparing models. A good candidate for coincident indicator is the gross non-farm domestic product, which is used as a proxy for the business cycle, whilst the industrial materials price is used as the leading indicator. The paper concludes with section 5.

2. The Model

The SSM is represented as follows

$$y_t = l_{t-1} + g_{t-1} + e_t \tag{1}$$

$$l_t = l_{t-1} + g_{t-1} + \alpha e_t \tag{2}$$

$$g_t = \mu_1 s_t + \mu_0 , \tag{3}$$

$$prob(s_t = 1 | s_{t-1} = 1) = p , \tag{4}$$

$$prob(s_t = 0 | s_{t-1} = 0) = q , \tag{5}$$

where y_t is the observed value, l_t represents the unobserved level at time t, g_t is the unobserved growth at time t, e_t's are independent and normally distributed disturbances with mean 0 and variance σ^2, α is the level smoothing parameter, p and q are the transition probabilities, s_t is the unobserved state of the system (or economy) at time t, assumed to follow a Markov model of order 1, that is

$$P(s_t | S_{t-1}, Y_t) = P(s_t | s_{t-1}) , \tag{6}$$

where $S_t = (s_0, s_1, \ldots, s_t)'$ and $Y_t = (y_1, y_2, \ldots, y_t)'$. The parameters μ_1 and μ_0 together define the two levels of growth. During an expansion, $s_t = 1$ and the growth rate is given by $g_t = \mu_1 + \mu_0$, whereas during a recession, $s_t = 0$ and the growth rate is given by $g_t = \mu_0$. As we require 'expansion' to have a higher growth rate than 'recession', we impose the constraint $\mu_1 > 0$.

By relaxing the conditions (4) and (5), the SSMVP model is represented by the equations (1), (2), (3) and

$$prob(s_t = 1 | s_{t-1} = 1) = p_t = H(z_t'\beta_1), \quad (4')$$

$$prob(s_t = 0 | s_{t-1} = 0) = q_t = H(z_t'\beta_0), \quad (5')$$

where z_t is a *m*-vector of known values, β_1 and β_0 are *m*-vectors of hyper parameters used to describe the transition probabilities p_t and q_t through the function *H*. Note here that if z_t in (4′) and (5′) is constant, all the transition probabilities p_t and q_t are equal respectively to fixed values *p* and *q*, hence the SSMVP will collapse to the SSM.

As this is essentially a simple latent variable model, the function *H* may be chosen as any known cumulative density function (*cdf*), which has the desirable consequence of forcing the computed transition probabilities to lie between zero and one. In this paper, *H* is chosen as the standard Gaussian *cdf*, with the results that the latent switching state probabilities are linked to the leading indicator using a probit model. Filardo and Gordon (1998) used this type of *cdf*, while the other *cdf* used in the literature is the logistic *cdf* (Filardo, 1994; Diebold et al, 1994).

The vector parameter λ to estimate is composed of three blocks of parameters. The first is λ_1, that is constituted from the smoothing parameter α, the variance of the errors σ^2 and the initial value of the state vector l_0. The second is λ_2 that is constituted from the two switching components μ_1 and μ_0, and the third is λ_3, that is constituted from the transition probabilities *p* and *q* in the SSM and the hyper parameters β_1 and β_0 in the SSMVP. Partitioning the parameter set into these three blocks is convenient because conditional on λ_3, the same algorithm is used to estimate λ_1 and λ_2 in the SSM and SSMVP. Moreover, λ_1 is associated with the linear structural model corresponding to exponential smoothing with a constant growth term and estimated in Forbes, Snyder and Shami (2000).

The likelihood function for the SSM model can be constructed from consideration of the joint probability distribution of the observed data and unobserved state variables. The joint probability of the observed data, $Y_n = (y_1, y_2, \ldots, y_n)'$, and the unobserved state vector, $S_{n-1} = (s_0, s_1, \ldots, s_{n-1})'$, given λ has the form

$$f(Y_n, S_{n-1}|\lambda) = f(y_1|s_0,\lambda) f(s_0|\lambda) \prod_{t=2}^{n} f(y_t|Y_{t-1}, S_{t-1}, \lambda) f(s_{t-1}|S_{t-2}, \lambda). \quad (7)$$

The likelihood function of the parameters, $L(\lambda|Y_n) = p(Y_n|\lambda)$, is calculated for any particular value of λ by averaging (7) over all possible 2^n values of S_{n-1}.

We will show how to compute the Bayesian posterior probability distribution in the next section utilising the special structure of the model. From the measurement equation in (1) and conditional on λ, S_{t-1} and Y_{t-1}, the distribution of y_t is normal with mean $l_{t-1} + g_{t-1}$ and variance σ^2. Let $\delta = 1 - \alpha$. By substituting the value of the noise term, $e_t = (y_t - l_{t-1} - g_{t-1})$, from the measurement equation (1) into the level transition equation (2) yields

$$l_t = \delta l_{t-1} + \delta g_{t-1} + \alpha y_t. \quad (8)$$

Back solving to time $t = 1$, and substituting into (1) we obtain

$$y_t = \delta^{t-1} l_0 + \sum_{j=1}^{t} \delta^{j-1} g_{t-j} + \sum_{j=1}^{t-1} \alpha \delta^{j-1} y_{t-j} + e_t, \quad (9)$$

where $g_0 = \mu_1 s_0 + \mu_0$. Note (9) can be conveniently rearranged as

$$\tilde{y}_t = \tilde{x}_t l_0 + e_t, \quad (10)$$

where

$$\tilde{y}_t = y_t - \sum_{j=1}^{t} \delta^{j-1} g_{t-j} - \sum_{j=1}^{t-1} \alpha \delta^{j-1} y_{t-j} \quad \text{and} \quad \tilde{x}_t = \delta^{t-1}. \quad (11)$$

Therefore, $f(y_t|Y_{t-1}, S_{t-1}, \lambda)$ can be computed in (7) using

$$f\left(y_t | Y_{t-1}, S_{t-1}, \lambda\right) \propto \sigma^{-1} \exp\left\{-\frac{1}{2\sigma^2}\left(\widetilde{y}_t - \widetilde{x}_t l_0\right)^2\right\}. \quad (12)$$

Note here that for given values of S_{n-1}, λ_2 and α, the linear regression estimates from (10) are

$$\hat{l}_0 = \left(\sum_{t=0}^{n-1} \delta^{2t}\right)^{-1} \sum_{t=1}^{n} \delta^{t-1} \widetilde{y}_t , \quad (13)$$

$$\hat{V}_{l_0} = \sigma^2 \left[\sum_{t=0}^{n-1} \delta^{2t}\right]^{-1}, \quad (14)$$

and $$SSE = \sum_{t=1}^{n} \left(\widetilde{y}_t - \delta^{t-1} \hat{l}_0\right)^2 . \quad (15)$$

3. A Bayesian Analysis

Taking the likelihood function, as discussed in Section 2, and the joint prior distribution $P(\lambda)$, we can construct the posterior distribution for the unknown parameters contained in λ using Bayes' theorem

$$P(\lambda | Y_n) = \frac{L(\lambda | Y_n) P(\lambda)}{P(Y_n)} . \quad (16)$$

However, direct Bayesian inference about λ in the SSM and consequently in the SSMVP is not available analytically. Hence, we use an MCMC technique (Gibbs sampling) that utilises the previously suggested partitioning of the parameter λ.

Sampling Method

The Gibbs procedure uses sampling from the following distributions

$$P(\lambda_1 | Y_n, S_{n-1}, \lambda_2, \lambda_3)$$

$$P(\lambda_2 | Y_n, S_{n-1}, \lambda_1, \lambda_3)$$

$$P(\lambda_3 | Y_n, S_{n-1}, \lambda_1, \lambda_2)$$

$$P(S_{n-1} | Y_n, \lambda_1, \lambda_2, \lambda_3).$$

If λ_3 is known, the procedure to find the conditional distributions $P(\lambda_1 | Y_n, S_{n-1}, \lambda_2, \lambda_3)$, $P(\lambda_2 | Y_n, S_{n-1}, \lambda_1, \lambda_3)$ and $P(S_{n-1} | Y_n, \lambda_1, \lambda_2, \lambda_3)$ is the same in the SSM and the SSMVP. In the case of the SSM we refer the reader to Shami and Forbes (2000) where the Gibbs procedure to estimate λ and S_{n-1} is detailed. The aim here is to extend this procedure to accommodate the estimation of the hyper parameters β_1 and β_0.

Prior Distributions

To complete a Bayesian analysis, a joint prior distribution for the unknown parameters must be specified. As parameters will be sampled in blocks, we specify the general form of the joint prior distribution by

$$P(\lambda) \propto P(\lambda_1)P(\lambda_2)P(\lambda_3). \tag{17}$$

We follow Forbes, Snyder and Shami (2000) by imposing

$$P(\lambda_1) = P(l_0, \alpha, \sigma^2) \propto \sigma^{-2} P(\alpha), \tag{18}$$

for $-\infty < l_0 < \infty$, $\sigma^2 > 0$ and $0 < \alpha < 2$. The limits of α are derived by writing the SSM as an ARIMA model. By taking the first difference of the level in (2) and substituting the result into the first difference of (1), we obtain an ARIMA(0,1,1) with drift. The moving average coefficient is equal to $\alpha - 1$. The invertibility condition of the ARIMA process leads us to impose that the absolute value of $\alpha - 1$ should be less than one, which translates into the constraint $0 < \alpha < 2$. As the algorithm we detail is not sensitive to the choice of $P(\alpha)$, we leave the notation general at this stage. In our example, we choose a uniform distribution.

The marginal priors for λ_2 and λ_3 are chosen to simplify the Gibbs sampling algorithm. We select flat prior distributions for λ_2 and λ_3, with zero mean and diagonal covariance matrices with large values. However, to account for the restriction imposed on the parameters μ_1 and μ_0, we choose

$$P(\lambda_2) = I_{\mu_1 > 0}, \tag{19}$$

so that observations at times corresponding to an expansion have a higher growth rate than those corresponding to a recession.

Conditional Distribution of λ_3

In this section, as we are really only interested in the conditional distribution $P(\lambda_3 | Y_n, S_{n-1}, \lambda_1, \lambda_2)$, we will present the extension to the SSM procedure which will cover the estimation of β_1 and β_0, and state the results concerning the other parameters (for more details, see Shami and Forbes; 2000). We will define the transition probabilities in terms of conditional latent probit variables u_t and use a modified version of Albert and Chib (1993).

Equations (4′) and (5′) can be described using the latent variable u_t,

$$\begin{cases} (u_t | s_{t-1} = 1) \sim N(z'_t \beta_1, 1) \\ (u_t | s_{t-1} = 0) \sim N(z'_t \beta_0, 1), \end{cases} \tag{20}$$

with $\begin{cases} s_t = 1 \text{ if } u_t > 0 \\ s_t = 0 \text{ if } u_t \leq 0, \end{cases}$

where z_t is the leading indicator series. Note here that the values of s_t and s_{t-1} impact on the distribution of $u_t | S_n, Y_n, \lambda$.

Given $\pi(\beta)$, the prior probability function of the hyper parameter $\beta = (\beta_1 \quad \beta_0)$, the joint posterior density of the unobservable β and $U_n = (u_1, u_2, \ldots, u_n)'$, given the switching state vector S_{n-1}, is given by

$$\begin{aligned} &P(\beta, U_n | S_{n-1}) \\ &\quad \propto \pi(\beta) \prod_{t=1}^{n} \{[I(u_t > 0) I(s_t = 1) + I(u_t \leq 0) I(s_t = 0)] I(s_{t-1} = 1) \phi(u_t - z'_t \beta_1) \\ &\qquad\qquad + [I(u_t > 0) I(s_t = 1) + I(u_t \leq 0) I(s_t = 0)] I(s_{t-1} = 0) \phi(u_t - z'_t \beta_0)\}, \end{aligned} \tag{21}$$

where $I(z>0)$ and $I(z\le 0)$ are the indicator functions that are equal to one if $z>0$ and $z\le 0$ respectively, and zero otherwise, and ϕ is the standard normal *pdf*. This joint distribution (21) is complicated in the sense it is difficult to normalise and sample from directly. However, the computation of the marginal posterior distribution of β conditional on U_n and the marginal posterior distribution of U_n conditional on β are of standard form. Assuming a priori independence of β_1 and β_0, so that the prior $\pi(\beta)$ can be described by the product of $\pi(\beta_1)$ and $\pi(\beta_0)$, from (21) the posterior distribution of β given U_n is given by

$$\begin{cases} P(\beta_1|U_n,S_{n-1}) \propto \pi(\beta_1)\prod_{t=1}^{n} I(s_{t-1}=1)\phi(u_t - z'_t\beta_1) \\ P(\beta_0|U_n,S_{n-1}) \propto \pi(\beta_0)\prod_{t=1}^{n} I(s_{t-1}=0)\phi(u_i - z'_t\beta_0) \end{cases}. \qquad (22)$$

The full conditional posterior density of β_1 given U_n and S_{n-1} is the usual posterior density for the regression parameter in the normal linear model $U_1 = Z_1\beta_1 + e_1$. Here U_1 is the l-vector of the latent variables u_t's corresponding to s_{t-1}=1, Z_1 is the matching (l,m) given values, $l < n$ and e_1 is distributed $N_l(0, I_l)$. Using standard linear model results, since the prior distribution of β_1 is diffuse, then

$$\beta_1|S_{n-1},U_n \text{ is distributed } N_m(\hat{\beta}_1,(Z_1'Z_1)^{-1}), \qquad (23)$$

where $\hat{\beta}_1 = (Z_1'Z_1)^{-1}Z_1'U_1$. In the case where β_1 is assigned the proper conjugate $N(\hat{\beta}_{10},\hat{B}_1^{-1})$ prior, then the posterior distribution of β_1 given U is given by $N_m(\tilde{\beta}_1,\tilde{B}_1^{-1})$, where $\tilde{B}_1 = \hat{B}_1 + Z_1'Z_1$ and $\tilde{\beta}_1 = \tilde{B}_1^{-1}(\hat{B}_1\hat{\beta}_{10} + Z_1'U_1)$.

Similarly, the full conditional posterior density of β_0 given U_n and S_{n-1} is given by the Gaussian distribution

$$N_m(\hat{\beta}_0,(Z_0'Z_0)^{-1}), \qquad (24)$$

where $\hat{\beta}_0 = (Z_0'Z_0)^{-1}Z_0'U_0$. In this case, U_0 is the j-vector of the latent variables u_t's corresponding to s_{t-1}=0, Z_0 is the matching (j,m) given values, $j = n\text{-}l$.

Also, from (21), the posterior distribution of U_n conditional on β has a simple form. The u_t's are independent and distributed from one of four truncated normal distributions $p(u_t|S_{n-1},\beta)$, according to the values of s_t and s_{t-1}.

$(u_t|S_{n-1},\beta)$ follows $N(z'_t\beta_1,1)\times I_{>0}$ if $s_t = 1$ and $s_{t-1} = 1$,

$N(z'_t\beta_1,1)\times I_{\le 0}$ if $s_t = 0$ and $s_{t-1} = 1$,

$N(z'_t\beta_0,1)\times I_{>0}$ if $s_t = 1$ and $s_{t-1} = 0$,

$N(z'_t\beta_0,1)\times I_{\le 0}$ if $s_t = 0$ and $s_{t-1} = 0$. (25)

Once the parameter β is generated, the two vectors of transition probabilities are evaluated. Let p_t and q_t be the probabilities of staying in expansion and in recession at time t respectively. p_t is given by

$$p_t = p(s_t = 1|s_{t-1} = 1) = p(u_t > 0|s_{t-1} = 1)$$
$$p_t = p(u_t - z_t'\beta_1 > -z_t'\beta_1)$$
$$p_t = p(e_t > -z_t'\beta_1)$$
$$p_t = 1 - \Phi(-z_t'\beta_1) = \Phi(z_t'\beta_1), \tag{26}$$

where *cdfn* is the cumulative distribution function of the standard normal distribution. Similarly, q_t is given by

$$q_t = \Phi(-z_t'\beta_0) = 1 - \Phi(z_t'\beta_0). \tag{27}$$

Generation Procedure

To begin the generation, initial values of μ_1, μ_0 and S_{n-1} are needed. Given that we dispose of b burn-in samples once convergence is obtained, the initial switching states are generated arbitrarily. Thus we take the simplest case where the switching states are generated from fixed transition probabilities p and q. To ensure that switching between the states occurs at least once, we require at least one s_t be equal to one and at least one s_t be equal to zero. Next, a value for β is computed using the least square estimate $(Z'Z)^{-1}Z'S$. In this case, β_1 is equal to $(Z_1'Z_1)^{-1}Z_1'1_L$ and $\beta_0 = 0$, where 1_L is the unit vector. Thereafter, the generation from the conditional distributions proceeds as follows:

- $\lambda_1 | Y_n, S_{n-1}, \lambda_2, \lambda_3$

   1. $p(\alpha | Y_n, S_{n-1}, \lambda_2) \propto \left(\sum_{t=0}^{n-1} \delta^{2t} \right)^{-½} SSE^{-(n-1)/2}$, where $0 < \alpha < 2$ and SSE given by (15).
   2. $\sigma^2 | \alpha, Y_n, S_{n-1}, \lambda_2$ ~ Inverted Gamma with mean $SSE/(n\text{-}3)$ and variance $2SSE^2/(n\text{-}3)^2(n\text{-}5)$.
   3. $l_0 | \sigma^2, \alpha, Y_n, S_{n-1}, \lambda_2$ ~ Normal with mean $\hat{l}_0$ and variance $\hat{V}_{l_0}$, calculated in (13) and (14).

Steps 2 and 3 imply that $l_0 | \alpha \sim Student$ distribution with n-1 degrees of freedom.

- $\lambda_2 | Y_n, S_{n-1}, \lambda_1, \lambda_3$

   4. $\mu_1, \mu_0 | \alpha, \sigma^2, l_0, Y_n, S_{n-1}$ ~ truncated Normal from $N(\hat{\mu}, \Sigma_\mu) I_{\mu_1 > 0}$, where $\hat{\mu} = (W'W)^{-1}W'V$ and $\Sigma_\mu = (W'W)^{-1}\sigma^2$, with $W' = (w_1, w_2, \cdots, w_n)$ and $V = (v_1, v_2, \cdots, v_n)$ in the conditional regression $v_t = w_t{}' \mu + e_t$ obtained by substituting g_{t-1} from (3) into the measurement equation (1), $y_t - l_{t-1} = \mu_1 s_{t-1} + \mu_0 + e_t$, where $v_t = y_t - l_{t-1}$ and $w_t = (s_{t-1}, 1)'$.

- $\lambda_3 | Y_n, S_{n-1}, \lambda_1, \lambda_2$

   5. $U_n | Z_n, S_{n-1}, \beta_1, \beta_2$ ~ n independent truncated Normal from (25).
   6. $\beta_1, \beta_2 | Z_n, S_{n-1}, U_n$ ~ two independent Normal from (23) and (24).

- $S_{n-1} | Y_n, \lambda_1, \lambda_2, \lambda_3$

   7. The entire vector of switching states is generated as a block from $P(S_{n-1} | Y_n, \lambda) = P(s_{n-1} | Y_n, \lambda) \prod_{t=0}^{n-2} P(s_t | Y_n, \lambda, s_{t+1})$ using a developed

algorithm of forward filtering and backward sampling (for full details of the algorithm, see Shami and Forbes; 2000)

Computing Posterior Marginal Switching Probabilities

Once the Gibbs sampler algorithm has converged and b burn-in values discarded, a sample of size r from the posterior distribution is available and estimates of numerous features of the posterior are available. Forbes, Snyder and Shami (2000) detail how to obtain forecast distributions for the linear model, and those calculations can be directly extended for the SSM and SSMVP models.

Of particular interest is the posterior marginal switching probabilities $P(s_t | Y_n)$, which can be computed using Rao-Blackwellised estimators (Gelfand and Smith, 1990) as follows

$$P(s_t = 1 | Y_n) = \frac{1}{r}\sum_{k=1}^{r} P(s_t = 1 | Y_n, \lambda^{(k)}), \qquad (28)$$

where $P(s_t = 1 | Y_n, \lambda^{(k)})$ is the $(t+1)^{th}$ element of the vector $P(S_{n-1} = 1 | Y_n, \lambda^{(k)})$ and $\lambda^{(k)}$ is the sampled parameter value for $k = 1, 2, ..., r$.

4. Application to Australian GDP

In this section the Australian business cycle will be analysed by modelling a component of the GDP as the coincident indicator and the industrial materials prices (IMP) as the leading indicator. Other leading components such as "overtime worked" and "number of housing approvals" (see Boehm and Moore, 1984 or Layton, 1997 for the components of an Australian leading index) are used in this model and showed no difference in the results obtained from using the IMP. The observations y_t are taken as the natural logarithm of quarterly real non-farm GDP in 1998-1999 prices multiplied by 100 for the period 1969/1 to 2000/4. The indicators z_t are the growth rates of the quarterly IMP for the same period as the GDP but in 1995 prices. The GDP data is obtained from the "National Accounts" and the IMP data is obtained from the "OECD main economic indicator", the two accounts can be found in the DX software database. Figure 1 shows the quarterly data of non-farm GDP (left scale in 10^3) and the IMP (right scale).

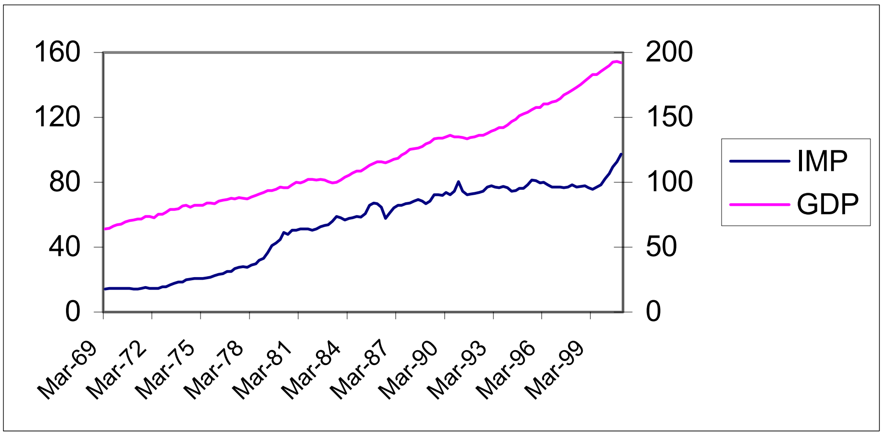

Figure 1: GDP and IMP observed data

For the estimation of the model we used the data from 1969/2 to 1998/4. We lost one quarter at the beginning of the data for use of the leading indicator and we left the last eight values for comparison purpose with the predicted values. The first 2000 iterations from the Gibbs sampling were burn-in values, to ensure that approximate convergence was obtained. An additional 3000 iterations were saved and used to draw inferences on the parameters and the switching states. To comment on the results obtained from our model, we will use the Bry-Boschan algorithm to find the peak and trough dates. It is a simple algorithm consisting of smoothing the data in a sequence of steps in the aim of distinguishing between real and spurious peaks and troughs. One condition is that the movement from a peak to a trough or from a trough to a peak cannot be shorter than two quarters. Another condition ensures that a complete cycle (peak to peak or trough to trough) must be at least five quarters long. The peaks and troughs dates from Bry-Boschan algorithm are given in Figure 2 (GDP growth) and Table1.

Peaks	2/75	2/77	2/82	4/85	2/90
Troughs	4/75	4/77	1/83	2/86	2/91

Table 1: Bry-Boschan dates

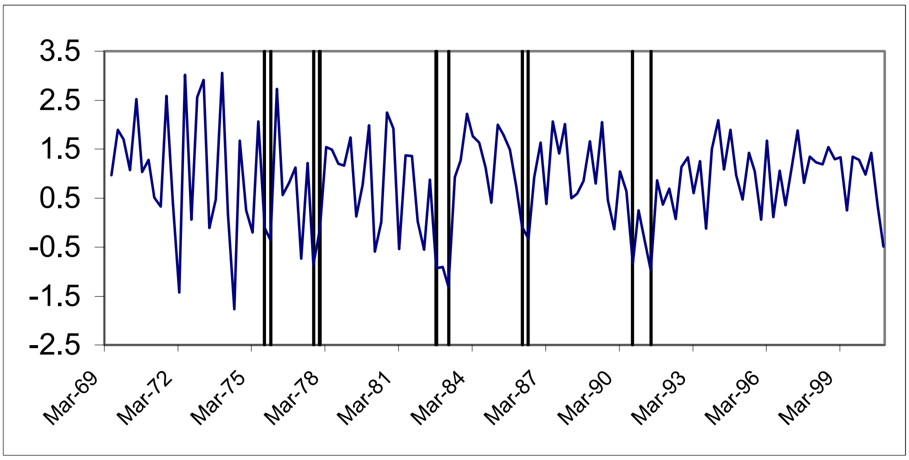

Figure 2: Bry-Boschan

The estimated posterior means from the SSMVP and the SSM are shown in Table 2 below along with the standard errors of the corresponding marginal posterior distributions.

	SSMVP		SSM	
	Estimate	SE	Estimate	SE
α	0.937	0.128	0.990	0.129
σ^2	0.769	0.172	0.865	0.178
l_0	1088.2	0.992	1088.3	0.996
μ_1	1.103	0.448	1.14	0.557
μ_0	-0.003	0.385	0.278	0.518
p			0.644	0.280
q			0.516	0.260
β_{10}	3.532	5.162		
β_{11}	-2.490	5.100		
β_{00}	3.806	9.182		
β_{01}	-3.700	9.042		

Table 2: Parameter estimates and their standard errors

The smoothing parameter value of 0.94 (0.99 in SSM) is close to one. The equivalent ARIMA(0,1,1) of the SSMVP (SSM) collapses to a near random walk with a switching drift. This is in line with previous studies suggesting the nonstationarity of GDP time series and the need to detrend the data before analysing it. While the growth during the expansion period (1.10 in SSMVP and 1.14 in SSM) is close to the observed average growth, the growth during the recessionary period μ_0 (-0.003 in SSMVP and 0.28 in SSM) is statistically insignificant. However, this insignificance does not change the nature of the switching behaviour of the series and this may be due to two effects. The first is the composition of the data, which has a few quarters of recession (13, see Table 1) divided into five periods out of 129 quarters. The second is the dynamic structure of the model that captures the change in the behaviour of the series. While the hyper parameters β_1 and β_0 also show insignificance in their values, the data concerned are still better estimated and predicted by incorporating the leading indicator.

Figure 3 shows the observations y_t (dark colour) and the estimated state vectors, which represent the levels l_{t-1} (light colour). Notice that the level l_t closely follows the observed data series, which is a main feature of a structural model.

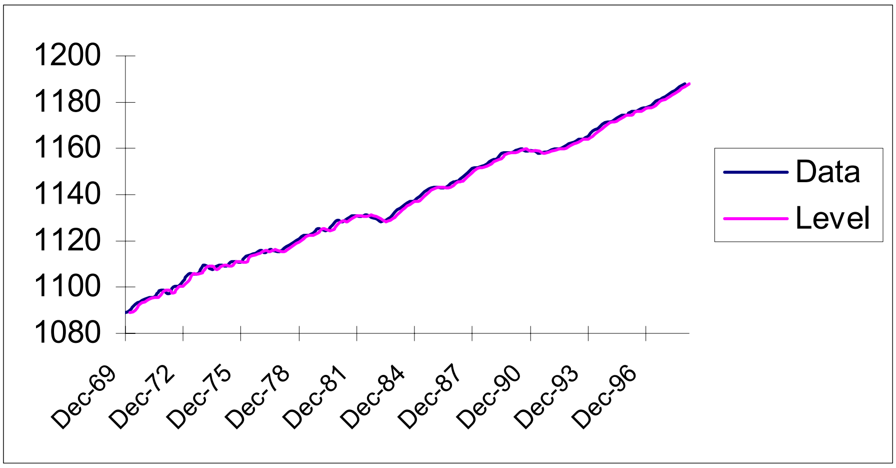

Figure 3: Observations and levels

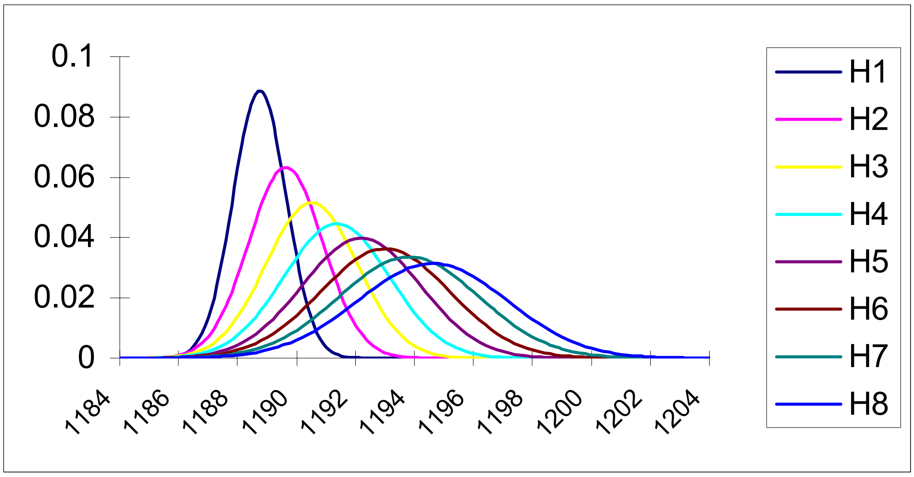

Figure 4: Marginal posterior distribution of the forecasts up to 8 horizons

Once the model is estimated, the parameters are used to forecast for 2 years ahead or 8 quarters up to the fourth quarter 2000. Rao-Blackwellised estimates are used to compute the mean and variance of the predictive values. The forecast distributions for all horizons are illustrated in Figure 4. These show how both the mean and the variance increase when the horizon increases. These distributions, which are a mixture of Gaussian distributions, appear symmetric. A similar figure is obtained for the SSM.

	Data	Mean	Std	L 95%	L 90%	U 90%	U 95%
Q1-1999	1189.18	1188.75	0.78	1186.88	1187.17	1190.13	1190.41
Q2-1999	1189.42	1189.64	1.52	1187.06	1187.46	1191.61	1192.01
Q3-1999	1190.76	1190.52	2.29	1187.38	1187.87	1192.96	1193.45
Q4-1999	1192.03	1191.37	3.07	1187.76	1188.32	1194.21	1194.77
Q1-2000	1193.00	1192.21	3.87	1188.18	1188.81	1195.40	1196.03
Q2-2000	1194.41	1193.01	4.69	1188.61	1189.30	1196.53	1197.21
Q3-2000	1194.76	1193.81	5.54	1189.07	1189.82	1197.62	1198.37
Q4-2000	1194.26	1194.60	6.43	1189.54	1190.34	1198.69	1199.49

Table 3 - Observations and estimates

Table 3 shows these predictions (Mean) along with the standard error of the corresponding posterior distributions (Std), 90% HPD intervals (L90% and

U90%) and 95% HPD intervals (L95% and U95%) where HPD means highest posterior density. They are also illustrated in Figure 5 with the 90% HPDs.

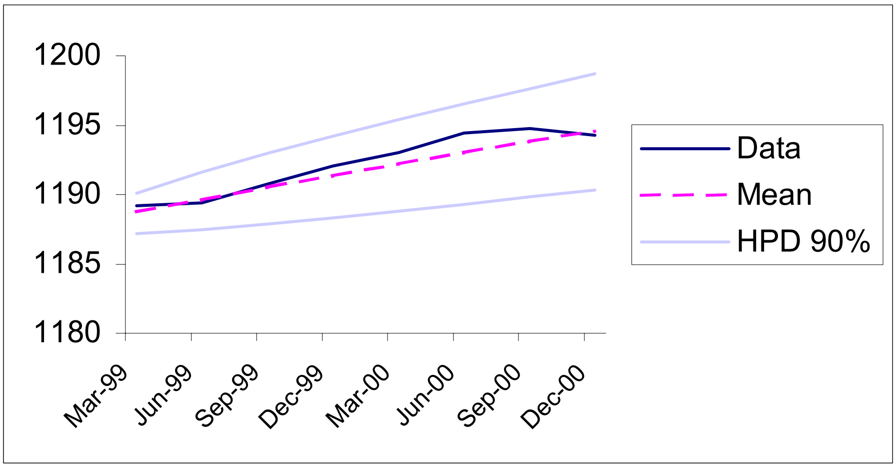

Figure 5: Observations, estimates and 90% interval estimates

A main feature of the SSMVP model is to capture the ups and downs in the series, and consequently show the expansion (s_t=1) and recession (s_t=0) phases of the business cycle by estimating the switching states.

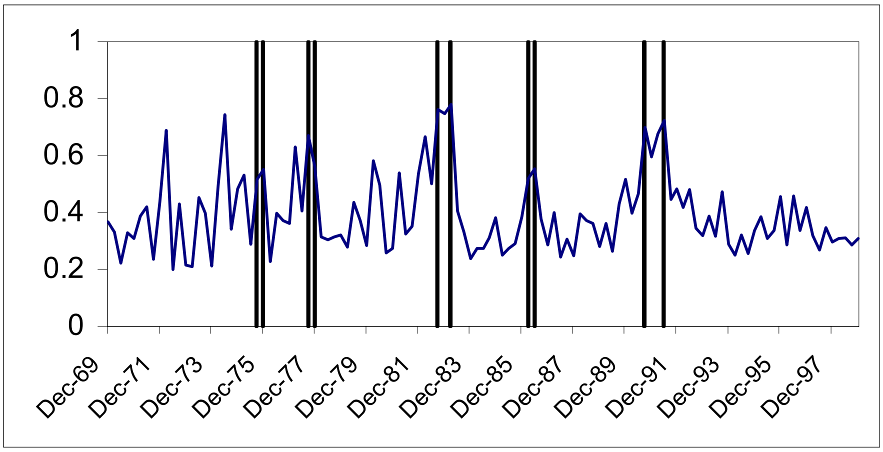

Figure 6: $P(s_t=0|Y_{t-1})$ - SSM

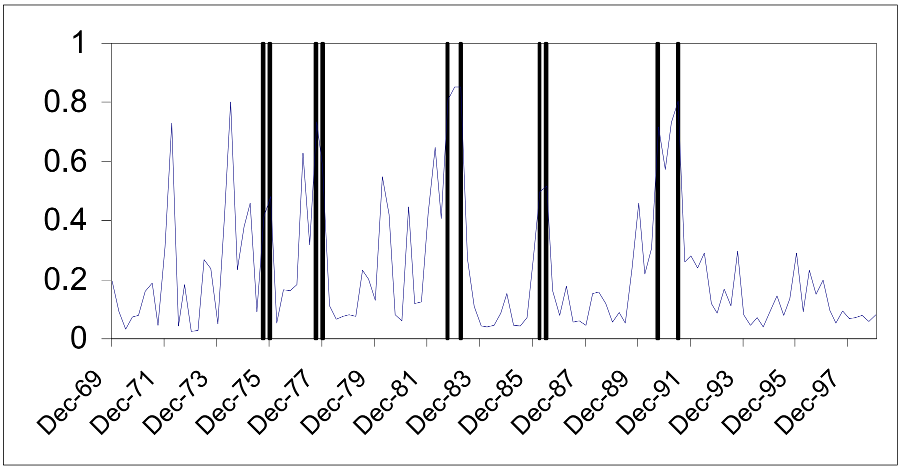

Figure 7: $P(s_t=0|Y_{t-1})$ - SSMVP

The two switching structural models prove very sensitive to the data. They capture every up and down in the movement of the GDP series considered here, which gives sometimes false alarms about the business cycle phases. However, if Bry-Boschan conditions are considered (at least two consecutive quarters of low growth to define a recession), the false signals showing only one quarter of low growth will be ignored as seen in Figure 6 (SSM) and Figure 7 (SSMVP) where the marginal filtered probability of being in low growth state, $P(s_t=0|Y_n)$ is presented with the dates of peak and troughs according to Bry-Boschan algorithm in vertical lines.

The signs of changing growths are similar in the two switching models, though it is more persistent in the SSMVP than the SSM. Comparing the results in Figures 6 and 7, when recessionary behaviour is captured by the models, we see that $P(s_t=0|Y_{t-1})$ has a higher value in the SSMVP than in the SSM and hence is closer to one, and when expansionary behaviour is present, $P(s_t=0|Y_{t-1})$ has a lower value in the SSMVP than in the SSM and hence is closer to zero.

Another way to compare different models is to evaluate the probability estimates. This can be done by many procedures. Here the two well-known measures that are described in Diebold and Rudebush (1989) are used. The first is the quadratic probability score (QPS) defined by Brier (1950) and given by the following

$$QPS = \frac{1}{N}\sum_{t=1}^{N} 2(p_{e,t} - p_{o,t})^2 , \qquad (29)$$

where $p_{e,t}$ is the estimated value of the probability at time t and $p_{o,t}$ is the observed value. The observed data used are the dates calculated by Bry-Boschan algorithm as contraction and expansion periods (see Figure 6). Like the usual mean squared error measure, the QPS provides a similar measure: a lower QPS implies that the prediction is more accurate. The other common measure is the log probability score (LPS), which is defined by

$$LPS = -\frac{1}{N}\sum_{t=1}^{N} [p_{o,t} \ln p_{e,t} + (1 - p_{o,t})\ln(1 - p_{e,t})]. \qquad (30)$$

Like QPS, a lower LPS implies that the prediction is more accurate. However, LPS penalises large mistakes more heavily than QPS, and while QPS is bounded by 0 and 2 ($0 < QPS < 2$), LPS has no upper bound ($0 < LPS < \infty$). Table 5 shows that the SSMVP outclasses the SSM in both the QPS and LPS measures.

	SSMVP	SSM
QPS	0.1315	0.2867
LPS	0.2488	0.4687

Table 4: Probability Scores

5. Conclusion

In this paper, the switching structural model (SSM) is presented and extended to include varying probabilities dependent on economic indicator (SSMVP). The Gibbs sampler used in Shami and Forbes (2000) to estimate the SSM is modified to accommodate the extension proposed using data augmentation methods. The two models were applied on quarterly Australian nonfarm GDP data by defining the expansion and recession phases of the business cycle as the two switching states. The extension is proved to be fruitful in that the inclusion of the economic indicator has helped to better estimate the probabilities of the two phases in the SSMVP than the SSM.

Also, by including the leading indicator into the structure of the SSMVP, the predictive power of the model is exploited whereas the Bry-Boschan procedure cannot be used for predicting the business cycle and the SSM is limited in its prediction ability. Such findings lead us to conclude that the SSMVP should be included in the family of business cycle models and that improved results can be obtained by using appropriate coincident and indicator indices instead of single indicators such the GDP and the IMP.

6. References

Albert, J. H. and Chib, S. (1993) Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts, *Journal of Business and Economic Statistics*, **11**, 1-15.

Aoki, M. (1987) *State Space Modelling of Time Series*, (Ed.) Springer-Verlag, Berlin.

Aoki, M. (1988) On alternative state space representations of time series models, *Journal of Economic Dynamics and Control*, **12**, 595-607.

Aoki, M. (1993) Interactions of real GNP business cycles in a three country time series model, *Journal of Forecasting*, **12**, 331-344.

Beaudry, P. and Koop, G. (1993) Do recessions permanently change output?, *Journal of Monetary Economics*, **31**, 149-163.

Boehm, E. A. and Moore, G. H. (1984) New economic indicators for Australia, 1949-84, *The Australian Economic Review*, **4-1984**, 34-56.

Boldin, M. D. (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. and Boschan, C. (1971) *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, (Ed.) National Bureau of Economic Research, New York.

Cecchetti, S. G., Lam, P. and Mark, N. C. (1990) Mean reversion in equilibrium asset prices, *The American Economic Review*, **80**, 398-418.

Clark, P. K. (1987) The cyclical component of U.S. economic activity, *Quarterly Journal of Economics*, **102**, 797-814.

Diebold, F. X., Lee, J. H. and Weinbach, G. C. (1994) Regime switching with time varying transition probabilities, in *Nonstationary Time-series Analysis and Cointegration*, (Ed.) C. Hargreaves, Oxford University Press, Oxford, pp.283-302.

Filardo, A. J. (1994) Business cycle phases and their transitional dynamics, *Journal of Business and Economic Statistics*, **12**, 299-308.

Filardo, A. J. and Gordon, S. F. (1998) Business cycle durations, *Journal of Econometrics*, **85**, 99-123.

Forbes,, C. S., Snyder, R. D and Shami, R. G. (2000) *Bayesian exponential smoothing*, Working paper 7/00, Department of Econometrics and Business Statistics, Monash University.

Gelfand, A. E. and Smith, A. F. M. (1990) Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association*, **85**, 398-409.

Hamilton, J. D. (1989) A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, **57**, 357-384.

Hamilton, J. D. (1990) Analysis of time series subject to changes in regime, *Journal of Econometrics*, **45**, 39-70.

Hamilton, J. D. and Lin, G. (1996) Stock market volatility and the business cycle, *Journal of Applied Econometrics*, **11**, 573-593.

Hansen, B. E. (1992) The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP, *Journal of Applied Econometrics*, **7**, S61-S82.

Harvey, A. C. (1984) A unified view of statistical forecasting procedures, *Journal of Forecasting*, **3**, 245-75.

Harvey, A. C. (1985) Trends and cycles in macroeconomic time series, *Journal of Business and Economic Statistics*, **3**, 216-227.

Kim, C.-J. (1994) Dynamic linear models with Markov switching, *Journal of Econometrics,* **60**, 1-22.

Kim, C.-J. and Nelson, C. (1998) Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching, *The Review of Economics and Statistics*, **80**, 188-201.

Krolzig, H-M. (2001) Business cycle measurement in the presence of structural change: International evidence, *International Journal of Forecasting*, **17**, 349-368.

Lam, P-S. (1990) The Hamilton model with a general autoregressive component, *Journal of Monetary Economics,* **26**, 409-432.

Layton, A. P. (1997) A new approach to dating and predicting Australian business cycle phase changes, *Applied Economics*, 29, 861-868.

Luginbuhl, R. and De Vos, A. (1999) Bayesian analysis of an unobserved component time series model of GDP with Markov switching and time varying growths, *Journal of Business and Economic Statistics*, **17**, 456-465.

Ord, J. K., Koehler, A. B. and Snyder, R. D. (1997) Estimation and prediction for a class of dynamic nonlinear statistical models, *Journal of the American Statistical Association*, **92**, 1621-29.

Potter, S. M. (1995) A nonlinear approach to US GNP, *Journal of Applied Econometrics*, **10**, 109-125.

Shami, R. G. and Forbes, C. S. (2000) *A structural time series model with Markov switching*, Working paper 10/00, Department of Econometrics and Business Statistics, Monash University.

Shami, R. G. and Snyder, R. D. (1998) *Exponential smoothing methods of forecasting and general ARMA time series representations*, Working paper 3/98, Department of Econometrics and Business Statistics, Monash University.

Snyder, R. D. (1985) Recursive estimation of dynamic linear models, *The Journal of the Royal Statistical Society - Series B*, **47**, 272-76.

Terasvirta, T. and Anderson, H. M. (1992) Characterizing nonlinearities in business cycles using smooth transition autoregressive models, *Journal of Applied Econometrics*, **73**, S119-S136.

Watson, M. W. (1986) Univariate detrending methods with stochastic trends, *Journal of Monetary Economics*, **18**, 49-75.