Censored Latent Effects Autoregression, with an Application to US Unemployment*

Philip Hans Franses†
Econometric Institute and RIBES
Erasmus University Rotterdam

and

Richard Paap‡
RIBES
Erasmus University Rotterdam

Econometric Institute Report 9841/A

Abstract

A new time series model is proposed to describe observed asymmetries in postwar unemployment data. We assume that recession periods, when unemployment increases rapidly, are caused by unobserved positive shocks. The generating mechanism of these latent shocks is a censored regression model, where linear combinations of lagged explanatory variables lead to positive shocks, while otherwise shocks are equal to zero. We apply our censored latent effects autoregression [CLEAR] to monthly US unemployment, where the positive shocks are found to depend on lagged oil prices, industrial production, the term structure of interest rates and a stock market index. The model fits the data well, and its out-of-sample forecasts appear to outperform those from alternative models.

*Discussions with Jaap van der Hart, James Stock and Jan Groen proved to be very helpful.

†Erasmus University Rotterdam, Econometric Institute, P.O. Box 1738, NL - 3000 DR Rotterdam, The Netherlands, e-mail: franses@few.eur.nl.

‡Corresponding author: Erasmus University Rotterdam, RIBES, P.O. Box 1738, NL - 3000 DR Rotterdam, The Netherlands, e-mail: paap@few.eur.nl.

1 Introduction

Postwar unemployment in many industrialized countries often displays the following two features: (i) in expansionary periods unemployment decreases slowly (perhaps toward a country-specific natural rate of unemployment), and (ii) in recessionary periods, which usually last much shorter than expansions do, the level of unemployment increases rapidly. The last feature leads to the empirical finding of a high degree of persistence in unemployment, which may be attributed solely to those large positive shocks during recessions, see Blanchard and Summers (1987) and Bianchi and Zoega (1998), among others.

The observed asymmetry in the time series pattern of unemployment established the motivation in several recent studies to propose modifications to linear models, which should be able to adequately describe the data. One class of these models is the (smooth) switching regime time series model, see Teräsvirta (1994), Granger and Teräsvirta (1993) and Tong (1990). This model allows for different dynamic structures across regimes, where the regimes are often defined by (changes in the) past observations of the time series of interest. For example, when the one-period lagged annual growth rate of unemployment exceeds a certain threshold level, the current value of the time series is said to correspond with a recession, see Skalin and Teraesvirta (1998) and Franses (1998), among others. The variable that determines those transitions can also be taken to be an exogenous variable like output growth, see Teräsvirta (1998).

An alternative class of models that can be useful to describe asymmetries in unemployment is the Markov switching model, initially put forward in Hamilton (1989), see also Hamilton (1994) for a survey. In contrast to the above models, the Markov model assumes that the transition between regimes is generated by a first order unobserved Markov process. Usually, the transition probabilities are assumed to be constant over time, but they may also depend on (lagged) exogenous variables, see Diebold *et al.* (1994) and Durland and McCurdy (1994). Recently, Bianchi and Zoega (1998) use a multi-state Markov switching model to construct a so-called shifting mean value model, where unemployment is assumed to undergo a sequence of level shifts.

In this paper we propose yet another time series model to describe the salient characteristics of postwar unemployment. The main idea behind our proposal is that we assume that recessionary periods for unemployment are caused by a few large positive shocks. In time series technology, we assume that recessionary periods correspond with a few large positive innovations. See Balke and Fomby (1994) for a related view on the time series behavior of macroeconomic aggregates. A second feature of our new model is that we do not use dummies for those large shocks as is done in Balke and Fomby (1994) but that we design a model that can generate those shocks as a function of lagged explanatory variables, like for example output, a term structure of interest rates and the oil price. In fact, our model may then be viewed as a leading indicator model for recessions for the series under interest. Since we wish to focus the attention solely to a few large positive shocks, we propose a censored model for the effects of the explanatory variables. As the shocks cannot be observed from the data, the latent effects of the explanatory variables have to be estimated. Finally, we model the dynamics of unemployment by an autoregressive structure, and therefore our model can best be called a censored latent effects autoregression [CLEAR] model.

The outline of our paper is as follows. In Section 2 we give the representation of the CLEAR model. We discuss parameter estimation, inference on unobserved variables and how one should calculate residuals (which in turn may be used for diagnostic purposes). In Section 3 we discuss how the CLEAR model differs from the time series models mentioned earlier. Without aiming to be complete, we also suggest modifications to the CLEAR model which may be useful for various alternative applications. In Section 4, we consider a CLEAR model for monthly postwar US unemployment. We compare forecasts from this model, and find improvement over other models. In Section 5 we conclude with some remarks.

2 The CLEAR Model

In this section we discuss various aspects of the CLEAR model, including representation, estimation, diagnostic checking using the estimated residuals, and forecasting.

2.1 Representation

A censored latent effects autoregressive model of order p [CLEAR(p)] for a univariate time series $\{y_t\}_{t=1}^T$ can be represented by

$$y_t = \mu + \sum_{i=1}^{p} \alpha_i y_{t-i} + v_t + \varepsilon_t, \tag{1}$$

where $\varepsilon_t \sim \text{NID}(0, \sigma_\varepsilon^2)$, $y_{p-1}, \ldots, y_1$ are fixed, and v_t is a censored latent variable defined by

$$v_t = \begin{cases} x_t'\beta + u_t & \text{if } x_t'\beta + u_t \geq 0 \\ 0 & \text{if } x_t'\beta + u_t < 0 \end{cases} \tag{2}$$

with $u_t \sim \text{NID}(0, \sigma_u^2)$, x_t a $(k \times 1)$ vector of exogenous variables including a constant and β an unknown $(k \times 1)$ parameter vector. As a positive value of v_t is added to the error term, the CLEAR(p) model contains an explicit description of an innovative outlier generating mechanism. The variable v_t is zero unless $x_t'\beta$ exceeds a stochastic threshold level $-u_t$, where u_t is a normal random variable. When the threshold is exceeded, v_t takes a positive value. Notice that if $\sigma_u = 0$ the model corresponds with a specific threshold model as we will indicate below in Section 3.1. By allowing $u_t \neq 0$, we introduce additional uncertainty as to whether linear combinations of exogenous variables, which may include leading indicator variables for the time series y_t, have a positive effect. Hence, the variance σ_u^2 is a measure for the quality of the predictive ability of the variables in x_t.

As v_t is a latent variable we can only make probability statements about its values. The probability that $v_t = 0$ equals the probability that $x_t'\beta + u_t < 0$, or more formally:

$$\Pr[v_t = 0 | x_t] = \int_{-\infty}^{-x_t'\beta} \frac{1}{\sigma_u} \phi(u_t/\sigma_u) du_t = \Phi\left(\frac{-x_t'\beta}{\sigma_u}\right), \tag{3}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function [pdf] and the cumulative density function [CDF] of the standard normal distribution, respectively. Likewise, the probability that $v_t \neq 0$ is given by

$$\Pr[v_t \neq 0 | x_t] = \int_{-x_t'\beta}^{\infty} \frac{1}{\sigma_u} \phi(u_t/\sigma_u) du_t = 1 - \Phi\left(\frac{-x_t'\beta}{\sigma_u}\right). \tag{4}$$

As v_t is a censored random variable, its expectation does not equal $x_t'\beta$. The expected value of the shock v_t follows from

$$\begin{aligned}
\mathrm{E}[v_t|x_t] &= \int_{-\infty}^{\infty} (x_t'\beta + u_t)\frac{1}{\sigma_u}\phi(u_t/\sigma_u)du_t \\
&= \int_{-x_t'\beta}^{\infty} (x_t'\beta + u_t)\frac{1}{\sigma_u}\phi(u_t/\sigma_u)du_t \\
&= \left(x_t'\beta + \frac{\sigma_u\phi(-x_t'\beta/\sigma_u)}{1-\Phi(-x_t'\beta/\sigma_u)}\right)\left(1-\Phi\left(\frac{-x_t'\beta}{\sigma_u}\right)\right),
\end{aligned} \tag{5}$$

where we use that if $u_t \sim \mathrm{N}(a, \sigma_u^2)$ then

$$\int_b^{\infty} \frac{u_t}{\sigma_u}\phi((u_t-a)/\sigma_u)du_t = \left(\frac{\sigma_u\phi((b-a)/\sigma_u)}{1-\Phi((b-a)/\sigma_u)} + a\right)\left(1-\Phi\left(\frac{b-a}{\sigma_u}\right)\right), \tag{6}$$

see Johnson and Kotz (1970, p. 81–83).

2.2 Estimation

The estimation of the model parameters $\theta = \{\mu, \alpha_1, \ldots, \alpha_p, \sigma_\varepsilon, \beta, \sigma_u\}$ can be done by maximum likelihood. To derive the likelihood function, we first consider the pdf of y_t given its past $Y_t = (y_{t-1}, \ldots, y_1)$ and given v_t, that is

$$f(y_t|Y_{t-1}, v_t; \theta) = \frac{1}{\sigma_\varepsilon}\phi\left(\frac{e_t - v_t}{\sigma_\varepsilon}\right), \tag{7}$$

where we write ε_t as the difference between $e_t = y_t - \mu - \sum_{i=1}^p \alpha_i y_{t-i}$ and v_t to denote that the pdf is conditional on v_t. The unconditional pdf of y_t given its past equals

$$\begin{aligned}
f(y_t|Y_{t-1}, x_t; \theta) = \Pr[v_t = 0|x_t] f(y_t|Y_{t-1}, v_t; \theta)|_{v_t=0} \\
+ \int_{-x_t'\beta}^{\infty} \frac{1}{\sigma_u}\phi\left(\frac{u_t}{\sigma_u}\right) f(y_t|Y_{t-1}, v_t; \theta)|_{v_t = x_t\beta + u_t} du_t,
\end{aligned} \tag{8}$$

where $\Pr[v_t = 0|x_t]$ is defined as in (3). Note that the model for y_t is in fact a continuous mixture with a discrete mixing distribution in $v_t = 0$ and a continuous mixing distribution for the remaining values of v_t. The log likelihood function equals the sum of the logarithms of the unconditional pdfs of y_t given in (8)

$$\mathcal{L}(Y_T|X_T; \theta) = \sum_{t=1}^{T} \ln(f(y_t|Y_{t-1}, x_t; \theta)), \tag{9}$$

where $X_T = (x_T, \ldots, x_1)$. This log likelihood function can easily be maximized using standard optimization algorithms, like Newton-Raphson. These algorithms can be based on numerical derivatives of the log likelihood function or the analytical derivatives shown in the appendix below. To decrease the computational burden in computing the value of the log likelihood function, the integral in (8) can be simplified using the following result

$$\begin{aligned}
&\int_b^\infty \frac{1}{\sigma_u}\phi(u_t/\sigma_u)\frac{1}{\sigma_\varepsilon}\phi((u_t-a)/\sigma_\varepsilon)du_t = \\
&\quad \int_b^\infty \frac{1}{\sqrt{2\pi}\sigma_u}\exp(-\frac{1}{2}(u_t/\sigma_u)^2)\frac{1}{\sqrt{2\pi}\sigma_\varepsilon}\exp(-\frac{1}{2}((u_t-a)/\sigma_\varepsilon)^2) = \\
&\quad \frac{1}{\sqrt{2\pi}\sigma_u\sigma_\varepsilon}\sigma\exp(-\frac{1}{2}(-\sigma^2\sigma_\varepsilon^{-4}+\sigma_\varepsilon^{-2})a^2)\int_b^\infty \frac{1}{\sigma}\phi((u_t-\sigma_\varepsilon^{-2}\sigma^2 a)/\sigma)du_t = \\
&\quad \frac{1}{\sqrt{2\pi}\sigma_u\sigma_\varepsilon}\sigma\exp(-\frac{1}{2}(-\sigma^2/\sigma_\varepsilon^2+1)a^2/\sigma_\varepsilon^2)(1-\Phi((b-\sigma_\varepsilon^{-2}\sigma^2 a)/\sigma))
\end{aligned} \tag{10}$$

where $\sigma^2 = \frac{\sigma_u^2\sigma_\varepsilon^2}{\sigma_u^2+\sigma_\varepsilon^2}$. Note that in our case $a = e_t - x_t'\beta$ and $b = -x_t'\beta$.

2.3 Inference on the Unobserved $\mathbf{v}_t$

The sequence $\{v_t\}_{t=1}^T$ is a sequence of unobserved stochastic variables. In Section 2.1 we have derived the unconditional probability that $v_t = 0$. This probability does not depend on the values of the time series y_t. However, if we know the values of Y_t, inference on v_t can be made conditional on Y_t. In fact, the probability that $v_t = 0$ given Y_t equals

$$\Pr[v_t = 0|Y_t, x_t] = \frac{\Pr[v_t = 0|x_t]f(y_t|Y_{t-1}, v_t;\theta)|_{v_t=0}}{f(y_t|Y_{t-1}, x_t;\theta)}, \tag{11}$$

where $\Pr[v_t = 0|x_t]$, $f(y_t|Y_{t-1}, v_t;\theta)$ and $f(y_t|Y_{t-1}, x_t;\theta)$ are defined in (3), (7) and (8), respectively. This conditional probability can be compared with the ex-post probabilities of the state variables in a Markov switching model, see *e.g.*, Hamilton (1989). Therefore, they may also be used to analyze turning points in the business cycle for the series under consideration.

Using the same arguments as above the expected value of the shock v_t given the values Y_t equals

$$\mathrm{E}[v_t|Y_t, x_t] = \frac{\int_{-x_t'\beta}^\infty (x_t'\beta+u_t)\frac{1}{\sigma_u}\,\phi\left(\frac{u_t}{\sigma_u}\right)f(y_t|Y_{t-1}, v_t;\theta)|_{v_t=x_t'\beta+u_t}du_t}{f(y_t|Y_{t-1}, x_t;\theta)}. \tag{12}$$

This expression can be simplified by combining the results in (10) and (6). The conditional expectation (12) gives an estimate of the magnitude of a shock at time t. For practical purposes, the quantities in (11) and (12) may be useful to calculate.

2.4 Residuals and Diagnostic Checking

The CLEAR model contains an unobserved component and hence replacing the parameters with their maximum likelihood values does not automatically lead to properly defined residuals. An easy way to construct residuals amounts to putting the error terms u_t and ε_t equal to zero. The estimated residuals denoted as $\hat{\varepsilon}_t$ would then be given by

$$\hat{\varepsilon}_t = y_t - \hat{\mu} - \sum_{i=1}^{p} \hat{\alpha}_i y_{t-1} - x_t'\hat{\beta}\, \mathbb{I}[x_t'\hat{\beta} \geq 0], \tag{13}$$

where $\mathbb{I}[\cdot]$ is an indicator function taking a value of 1 if the argument is true and zero elsewhere.

This way of generating residuals is somewhat naive since we already have seen that $\mathrm{E}[v_t|x_t]$ does not equal $x_t'\beta$[1]. A better approach to estimate residuals is to take the expectation of ε_t given Y_t

$$\begin{aligned} \hat{\varepsilon}_t &= \mathrm{E}[\varepsilon_t|Y_t, x_t] \\ &= \mathrm{E}[y_t - \hat{\mu} - \sum_{i=1}^{p} \hat{\alpha}_i y_{t-i} - v_t|Y_t, x_t] \\ &= y_t - \hat{\mu} - \sum_{i=1}^{p} \hat{\alpha}_i y_{t-i} - \mathrm{E}[v_t|Y_t, x_t], \end{aligned} \tag{14}$$

where $\mathrm{E}[v_t|Y_t, x_t]$ is defined in (12). Hence the corresponding fit of the model is simply $\hat{\mu} + \sum_{i=1}^{p} \hat{\alpha}_i y_{t-i} + \mathrm{E}[v_t|Y_t, x_t]$.

Finally, one may calculate the one-step ahead forecast errors and treat these as residuals. A residual at time t is then defined as the difference between y_t and $\mathrm{E}[y_t|Y_{t-1}, x_t]$. The expectation of y_t given its past is simply

$$\mathrm{E}[y_t|Y_{t-1}, x_t] = \int_{-\infty}^{\infty} f(y_t|Y_{t-1}, x_t; \theta) dy_t, \tag{15}$$

[1]The residuals defined in (13) are in fact the residuals of the closely related threshold model defined in (22) below.

where $f(y_t|Y_{t-1}, x_t; \theta)$ is given in (8). Note that we need numerical integration methods to evaluate this integral. The one-step ahead forecast residuals now equal

$$\hat{\varepsilon}_t = \mathrm{E}[y_t|Y_{t-1}, x_t] - y_t \tag{16}$$

The estimated residuals may be used for diagnostic checking. To test for neglected serial correlation in the residuals we propose the auxiliary regression

$$\hat{\varepsilon}_t = \omega + \rho\hat{\varepsilon}_{t-1} + \eta_t. \tag{17}$$

The nulhypothesis $\rho = 0$, *i.e.*, the absence of first order serial correlation can be tested using an F-test. Likewise, tests for higher order serial correlation can be constructed. The presence of first order ARCH effects can be tested via the regression

$$\hat{\varepsilon}_t^2 = \kappa + \psi\hat{\varepsilon}_{t-1}^2 + \xi_t. \tag{18}$$

The nulhypothesis $\psi = 0$ corresponds to the absence of first order ARCH effects. This hypothesis can again be tested with an F-test. Tests for higher order ARCH effects proceed in the same way. Finally, we can use the $\chi^2(2)$ test for normality of the residuals.

2.5 Forecasting

In the previous subsection we have discussed the in-sample fit of the CLEAR model. We have shown that it is possible to use one-step ahead forecasts to generate residuals. The one-step ahead forecast at time T conditional on Y_T and x_{T+1} is given by

$$\mathrm{E}[y_{T+1}|Y_T, x_{T+1}] = \int_{-\infty}^{\infty} f(y_{T+1}|Y_T, x_{T+1}; \theta) dy_{T+1}, \tag{19}$$

where $f(y_{T+1}|Y_T, x_{T+1}; \theta)$ is defined as in (8). Notice that we assume knowledge of the values of explanatory variables in the innovation outlier model at time $T+1$. Likewise, we can construct two-step ahead forecasts. The two-step ahead forecast conditional on (x_{T+2}, x_{T+1}) and Y_T is given by

$$\mathrm{E}[y_{T+2}|Y_T, x_{T+2}, x_{T+1}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y_{T+2}|Y_{T+1}, x_{T+2}; \theta)\ f(y_{T+1}|Y_T, x_{T+1}; \theta)\ dy_{T+2}\ dy_{T+1}. \tag{20}$$

Multistep ahead forecasts are defined in a similar way. The h-step ahead forecast, conditional on $(x_{T+h},\ldots,x_{T+1})$ and Y_T, equals

$$\begin{aligned}&\mathrm{E}[y_{T+h}|Y_T,x_{T+h},\ldots,x_{T+1}]=\\&\qquad\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(y_{T+h}|Y_{T+h-1},x_{T+h};\theta)\cdots f(y_{T+1}|Y_T,x_{T+1};\theta)\;dy_{T+h}\cdots dy_{T+1}.\end{aligned}\tag{21}$$

Since it may be impossible to evaluate the multiple integrals using numerical integration, we may have to use straightforward simulation techniques to evaluate the integrals.

The h-step ahead forecast (21) is conditional on the value of $(x_{T+h},\ldots,x_{T+1})$. If the vector x_t consists of lagged explanatory variables some of these values may be known at time T. However, for multi-step ahead forecasts it is likely that some future values of x are unknown at time T and that they have to be replaced by forecasts.

3 Comparison with Other Models and Extensions

In this section we indicate the links between our CLEAR model and existing models, without aiming to give a comprehensive account. Similarly, we sketch some potentially useful extensions.

3.1 Relation with Threshold and Mixture Models

Straightforward simplifications of the CLEAR model result in familiar threshold and mixture models. If we set $u_t=0$ we obtain the following threshold autoregressive [TAR] model

$$y_t=\begin{cases}\mu+\sum_{i=1}^{p}\alpha_i y_{t-i}+x_t'\beta+\varepsilon_t & \text{if } x_t'\beta\geq 0\\ \mu+\sum_{i=1}^{p}\alpha_i y_{t-i}+\varepsilon_t & \text{if } x_t'\beta<0.\end{cases}\tag{22}$$

This models differs from the CLEAR model in several ways. First, the exogenous variables x_t enter the model without error term u_t. Second, in the TAR model it is assumed to known at time t whether $x_t'\beta$ enters the AR model or not. In the CLEAR model we only know at time t the probability that $v_t\neq 0$. In a sense, the CLEAR model in (1) and (2) is therefore less restrictive than the TAR model (22). Identification and estimation (especially of β) of a threshold model as in (22) is very complicated, since regime switching is based on an unknown linear combination of variables, see also Chen (1995). Therefore, it may be more convenient to consider the smooth transition version of (22)

$$y_t = \mu + \sum_{i=1}^{p} \alpha_i y_{t-i} + x_t'\beta F(x_t'\beta) + \varepsilon_t \tag{23}$$

where F is for instance a logistic function, see also Granger and Teräsvirta (1993) and Teräsvirta (1998).

If u_t does not enter the AR part of the CLEAR model, we obtain the following model

$$y_t = \begin{cases} \mu + \sum_{i=1}^{p} \alpha_i y_{t-i} + x_t'\beta + \varepsilon_t & \text{if } x_t'\beta + u_t \geq 0 \\ \mu + \sum_{i=1}^{p} \alpha_i y_{t-i} + \varepsilon_t & \text{if } x_t'\beta + u_t < 0 \end{cases} \tag{24}$$

Like in the CLEAR model, the probability that $x_t'\beta$ enters the AR model in (24) is equal to the probability that $u_t \geq -x_t'\beta$. The model in (24) amounts to a discrete mixture of two AR models with mixing proportion $\Phi(-x_t'\beta/\sigma_u)$, see Everitt and Hand (1981) for a discussion about mixtures. Note that this mixing proportion changes over time as it depends on the value of x_t. Therefore, also the probability of being in a recession changes over time. The closest link of a CLEAR model to a Markov switching model is to allow the transition probabilities in the latter model, which determine the changes between the two regimes, to depend on $x_t'\beta$, as in Diebold *et al.* (1994).

3.2 Extensions

The CLEAR model (1) and (2) can be extended in several directions. A straightforward extension is to add moving error terms to (1) resulting in a CLEARMA model. Another extension of the model concerns the impact of the shock v_t. The shock v_t enters the AR(p) model (1) as an innovation outlier, which means that a positive v_t lifts the time series to a higher level after which it may return slowly to its equilibrium level (depending on the largest AR root). Note that new shocks may lift up the level of the series again before it reaches its equilibrium level. Naturally, we may also add v_t as an additive shock

$$(y_t - v_t) = \mu + \sum_{i=1}^{p} \alpha_i (y_{t-i} - v_{t-i}) + \varepsilon_t. \tag{25}$$

A positive shock v_t lifts the time series to a higher level but it immediately returns to its original level in the next period, unless of course $v_{t+1} \neq 0$.

Finally, it seems that our univariate CLEAR model can be easily extended to multivariate time series. Also, we then have opportunities to investigate if there is a common innovation outlier generating mechanism across variables.

4 US Unemployment

To illustrate our CLEAR model, we consider an application to seasonally adjusted monthly observed unemployment rate of the United States [US], 1969.01–1997.12. Figure 1 shows a plot of the log of the unemployment rate. We notice short periods characterized by large increases in unemployment, which can be called recessions, and longer periods with slow decline in the unemployment rate, the expansions. To model this behavior we consider the CLEAR model as in (1) and (2). We examine if the large increases during recessions are captured by the censored latent effect variable v_t. As explanatory variables for the censored latent effects, we use monthly seasonally adjusted US industrial production, the oil price in dollars deflated by seasonally adjusted US CPI, the Dow Jones index and the difference between the 10 year treasury with constant maturity and a 3-month treasury bill rate of the United States[2]. The last variable is known to be a good predictor for turning points, see for instance Harvey (1988), Estrella and Hardouvelis (1991) and Estrella and Mishkin (1997).

The estimation period is 1969.01–1997.12. Earlier observations are used as starting values. Denote y_t as the log of the unemployment rate. As explanatory variables we use o_t, the log of the real oil price, r_t, the difference between the long term and short term interest rate, d_t, the log of the Dow Jones index and i_t, industrial production. To make the time series of the explanatory variables approximately stationary we take first differences of the oil price, the Dow Jones index and industrial production. A lag order of $p = 1$

[2] The data are obtained from the internet site of the Federal Reserve Bank of St. Louis except for the Dow Jones index, which is taken from Datastream.

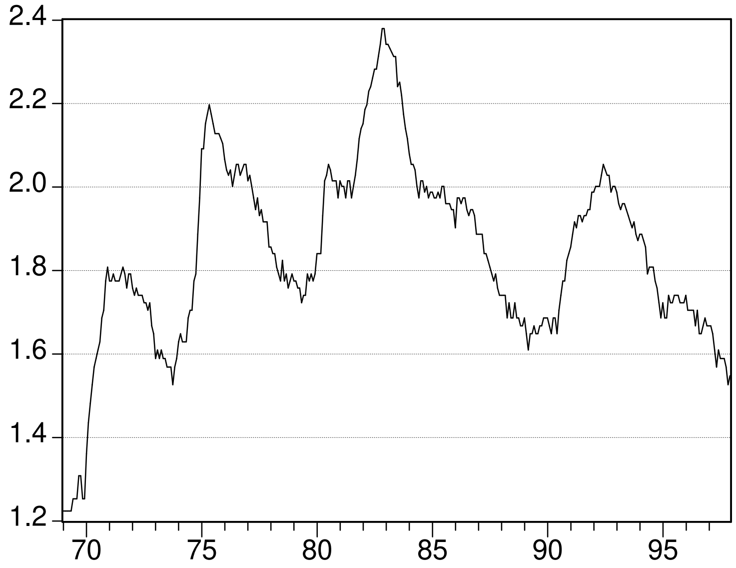

Figure 1: The logarithm of US unemployment rate, 1969.01–1997.12

turns out to be appropriate to capture the dynamics in (1). After some experimentation[3] we arrive at

$$y_t = \underset{(0.014)}{0.017} + \underset{(0.007)}{0.987}\, y_{t-1} + v_t + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}\,(0, \underset{(0.001)}{0.023^2}) \tag{26}$$

with

$$\begin{aligned} v_t = {} & \underset{(0.008)}{0.013} - \underset{(0.007)}{0.011}\,\Delta i_{t-2} - \underset{(0.005)}{0.017}\, r_{t-10} + \underset{(0.037)}{0.074}\,\Delta o_{t-12} - \\ & \underset{(0.087)}{0.243}\,\Delta d_{t-7} + u_t, \qquad u_t \sim \text{NID}\,(0, \underset{(0.006)}{0.026^2}), \end{aligned} \tag{27}$$

where estimated standard errors are given in parentheses. The large AR(1) coefficient in (26) indicates a very slow decay towards an equilibrium. Note that it is not necessary that this equilibrium is ever reached since positive shocks v_t may move the unemployment rate away from this equilibrium.

[3] We experimented with alternative lag structures in (27), and we found that in only one case (Δi_{t-7} instead of Δi_{t-2}) the optimal value of the log likelihood is slightly larger. However, in that case residual diagnostic tests indicate misspecification.

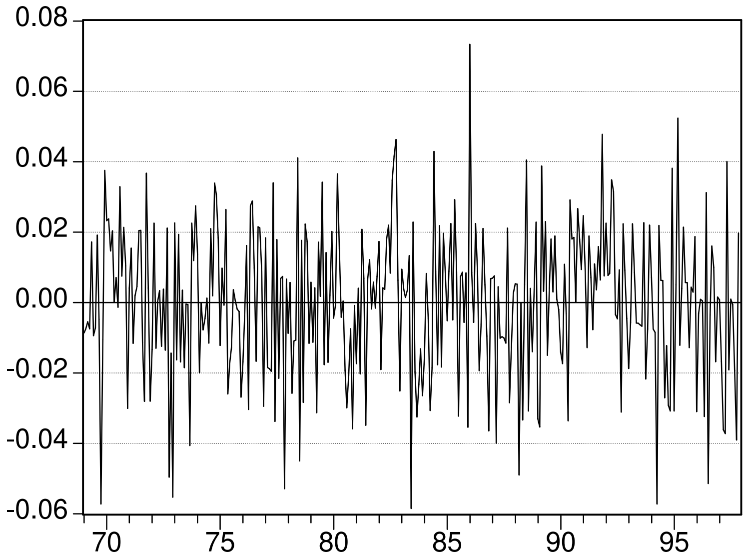

Figure 2: Residuals of the CLEAR model of the log of US unemployment rate

The coefficients of the explanatory variables in the censored regression (27) have the expected sign. Negative growth in industrial production, declining Dow Jones returns and a negative difference between long and short term interest rates increase the probability of a positive v_t in (26) and hence to a sudden increase in the level of unemployment. The same applies to an increase in the real oil price. The model suggests that negative growth in industrial production effects unemployment after 2 months. For the term structure, the oil price and the Dow Jones index, we find that their influence become apparent after 10, 12 and 7 months, respectively.

To analyze the fit of the model we compute the expected value of ε_t, $t = 1, \ldots, T$ given the data Y_T as proposed in (14). Figure 2 shows a plot of these residuals. An F-statistic for first order serial correlation based on regression (17) equals 2.85 which is not significant at the 5 % level of significance[4]. The F-test statistic for $\psi = 0$ in (18) to test for first order ARCH effects equals 0.18 with a p-value 0.67. Hence there seem no serious

[4]The same test statistics based on the naive residuals in (13) and on the one-step ahead forecast errors in (16) clearly reject serial correlation. The p-values are 0.98 and 0.79 respectively.

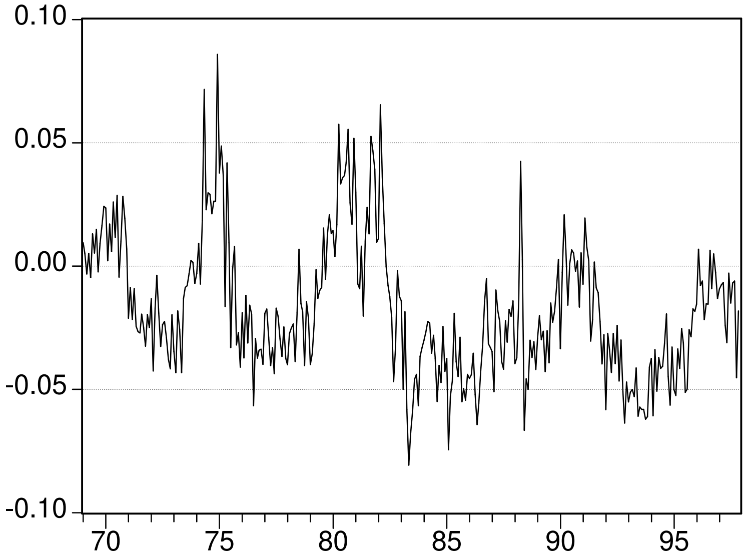

Figure 3: Fitted values of the censored latent regression ($x_t'\hat{\beta}$)

first order ARCH effects in the residuals. The $\chi^2(2)$ normality test on the residuals equals 0.18 and hence normality cannot be rejected. In sum, the CLEAR model in (26) and (27) does not seem to be misspecified.

Figure 3 shows a plot of the fitted values of the censored regression (27), that is $x_t'\hat{\beta}$. If we compare this graph with Figure 1 we see that the estimated linear combination of the explanatory variables follows the swings in the unemployment series quite well. Note again that the linear combination only enters the AR model (26) if $v_t \neq 0$. Figure 4 displays the conditional probabilities $\Pr[v_t \neq 0|Y_t, x_t]$. Likewise a Markov switching model, see for instance Hamilton (1989), we can interpret these conditional probabilities in terms of a business cycle. We may define a recession by 6 consecutive months for which $\Pr[v_t \neq 0 \mid Y_t, x_t] > 0.5$. A peak is defined by the last expansion observation before a recession. A trough is defined by the last observation in a recession. Table 1 shows the peaks and troughs together with the official NBER turning points. Our first observation from this table is that the conditional probabilities do not indicate a recession at the beginning of the 1990s. As can be seen from Figure 4, however, there is no sequence of six months in

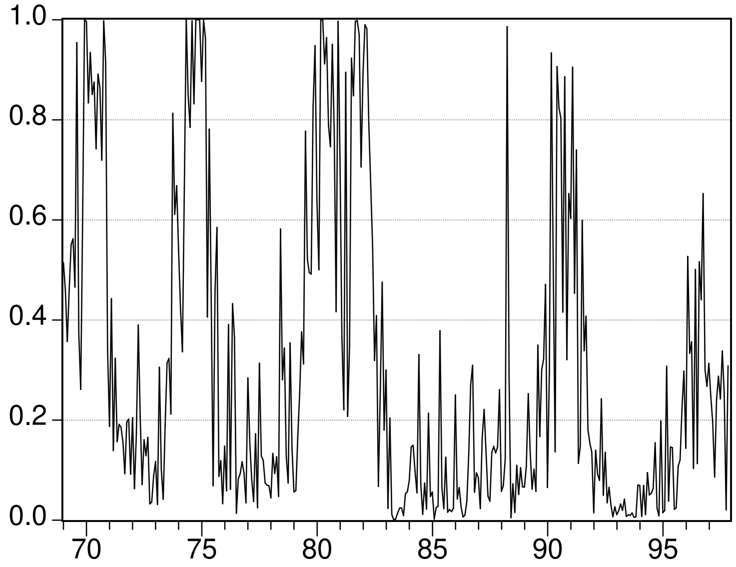

Figure 4: Conditional probabilities $\Pr[v_t \neq 0|Y_t, x_t]$

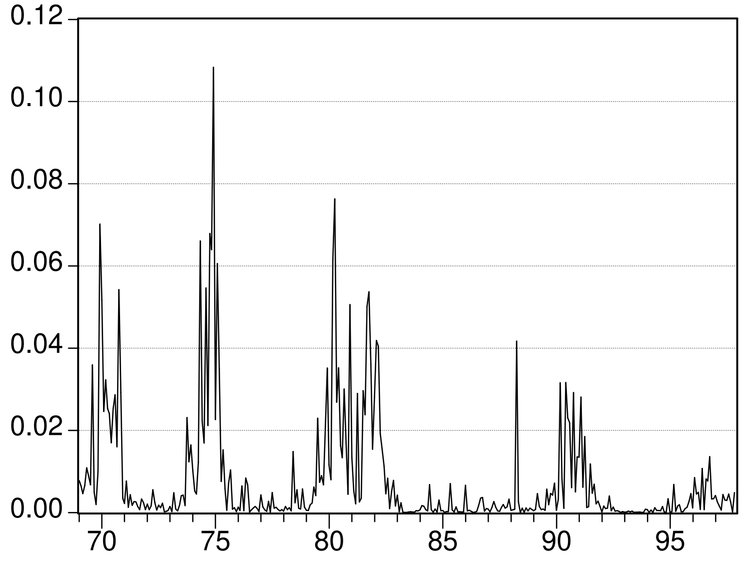

Figure 5: Conditional expectation of the censored latent effect $\mathrm{E}[v_t|Y_t, x_t]$

Table 1: Peaks and troughs for US unemployment based on conditional probabilities.

Unemployment[1]		NBER	
peak	trough	peak	trough
69.11	70.12	69.12	70.11
74.04	75.06	73.11	75.03
79.11	80.11	80.01	80.07
81.07	82.07	81.07	82.11
–	–	90.07	91.03

[1] A recession is defined by 6 consecutive months for which $\Pr[v_t \neq 0 \mid Y_t, x_t] > 0.5$. A peak corresponds with the last expansion observation before a recession and a trough with the last observation in a recession.

a row for 1990.04–1991.03 for which $\Pr[v_t \neq 0|Y_t, x_t] > 0.5$. Secondly, also the effects of the first oil crisis starts 5 months later at least in the case of the CLEAR model. If we take a closer look at Figure 4, we see that the conditional probabilities $\Pr[v_t \neq 0|Y_t, x_t]$ exceed 0.5 for 1973.11–1974.02, and only for 1974.03 this probability is below 0.5, causing the suggested recession to start as late as 1974.04.

In sum, the empirical evidence sofar suggests that the CLEAR model can potentially be used to actually understand which (combinations of) variables generate recessions.

Figure 5 shows the estimated conditional expectation of v_t, $\mathrm{E}[v_t|Y_t, x_t]$. The value of this expectation concerns the magnitude of the innovative shocks, which lift the unemployment rate to a higher level. Therefore, it can be used as an indication of the magnitude of a recession. The second recession in our sample (around 1974/75) seems to be the most serious recession, as it corresponds with the highest expected values of v_t. Apparently, the recession at the beginning of the 1990s did not have such a large impact on the unemployment rate level.

Finally, we compare the out-of-sample forecasting performance of the CLEAR model

Table 2: Forecast performance evaluation of the CLEAR model.

	criteria		sign test[1]		encompassing tests[2]	
model	RMSE	MAPE	fraction	p-value	I	II
			forecasting sample 1997.01–1997.12			
CLEAR	0.027	0.014				
AR(1)	0.029	0.015	10/12	(0.01)	0.51 (0.49)	2.37 (0.15)
ARX(1)	0.028	0.015	9/12	(0.02)	0.91 (0.36)	2.41 (0.15)
			forecasting sample 1996.01–1997.12			
CLEAR	0.027	0.013				
AR(1)	0.029	0.014	17/24	(0.01)	0.77 (0.39)	3.75 (0.07)
ARX(1)	0.029	0.015	17/24	(0.01)	1.80 (0.19)	5.35 (0.03)

[1] A nonparametric sign test. The first column displays the fraction that the forecast error of the CLEAR model is smaller in absolute value than the forecast error of the AR(1) and ARX(1) model. The second column displays the p-value for the corresponding nonparametric test for equal forecast accuracy.

[2] In column I we report encompassing tests (with p-values between parentheses) to test whether forecasts generated by the CLEAR model encompass forecasts generated by the AR(1) and ARX(1) model, while in column II we investigate whether the AR(1) and ARX(1) model forecasts encompass the CLEAR model forecasts.

with two closely related models. As competing models we take a simple AR(1) model and an ARX(1) model which is the CLEAR model without censoring and with imposing $u_t = 0$. In the ARX model it is assumed that the lagged explanatory variables enter the autoregression in all months, while in the CLEAR model only in some months. First, we hold out the last 12 observations and in a second exercise the last 24 observations from our sample and re-estimate the parameters of the CLEAR model. Next, we generate 12 (24) one-step ahead forecasts from the estimated CLEAR model using (19), and from the estimated AR(1) and ARX(1) models for the period 1997.01–1997.12 (and 1996.01–1997.12). Note that we can condition on x_t to compute the one-step ahead forecasts, as the models only contain lagged and not current values of the explanatory variables. The second and third columns of Table 2 show the root mean squared forecast error [RMSE] and the mean absolute percentage errors [MAPE] for the three models. The RMSE and MAPE for the forecasts of the CLEAR model are smaller than for the other two models, although the difference appear not to be very large.

The fourth and fifth columns present the outcomes of a nonparametric sign test for equal forecast accuracy. If we look at the forecasting sample 1997.01–1997.12, the forecast errors of the CLEAR model are 10 (9) times smaller in absolute value than the forecast error generated by the AR (ARX) model. Compared with the CDF of a binomial distribution with $p = 0.5$ and $n = 12$, these numbers are large enough to reject equal forecast accuracy at the 5% level. For the forecasting sample 1996.01–1997.12, the CLEAR model produces 17 times out of 24 smaller forecast errors in absolute value, which is again significant at the 5% level.

The final two columns of Table 2 show the outcomes of forecast encompassing tests. Let f_{T+h} be the forecast from the CLEAR model and $\bar{f}_{T+h}$ the forecast from a competing model. Then, the forecasts from the CLEAR model are said to encompass the forecasts from the competing model if the γ coefficient in the following regression model is zero

$$y_{T+h} - \bar{f}_{T+h} = \gamma(f_{T+h} - \bar{f}_{T+h}) + e_{T+h}, \text{ for } h = 1, \ldots, H, \tag{28}$$

where y_{T+h} is the true value, see Clements and Hendry (1993). To test for $\gamma = 0$ we use an F-statistic. The p-values reported in the final two columns show that the forecasts of the CLEAR model encompass the forecasts of the other two models. If we test whether the forecasts of the AR or ARX model encompass the forecasts of the CLEAR model we reject at the 15% level for the period 1997.01–1997.12 and at the 7% level or lower for the period 1996.01–1997.12.

To summarize this empirical section, the CLEAR model appears to describe the salient characteristics of the US unemployment rate rather well. The censored regression generates the positive shocks to unemployment during recession, and hence the included explanatory variables seem important from a forecasting point of view. Turning points based on the conditional probabilities for a positive shock match the peaks and troughs of the NBER reasonably well. Furthermore, the CLEAR model produces better out-of-sample forecasts than an AR(1) and an ARX(1) model for the period 1996.01–1997.12.

5 Concluding Remarks

In this paper we have proposed a new time series model to describe asymmetries in macroeconomic time series. The model consists of an autoregressive component and a component corresponding to positive innovations. These innovations are generated by a censored regression model, which contains (lagged) explanatory variables to describe the value of the shock. The innovative shocks only enter the time series model if their value exceeds a stochastic threshold level. As the exact values of the shocks are unknown and censored, our new time series model is called a censored latent effects autoregression [CLEAR].

The CLEAR model was illustrated for seasonally adjusted monthly US unemployment rate. The censored regression model, which generates the innovative shocks during recessions contained lagged values of leading indicators for unemployment, including the real oil price, a term structure, the Dow Jones index and industrial production. The models fitted the data well and its out-of-sample forecasts outperformed forecasts from alternative models.

The censored latent effects model provides ample possibilities for future research work. A new formal model identification strategy based on formal diagnostic tests, like Lagrange multiplier test for serial correlation and ARCH effects may be required. Extensions to moving average disturbances and additive shocks instead of innovative shocks are straightforward but may require more complicated algorithms to compute the likelihood function. The censored latent regression model can also be used to describe patterns in return volatility, as volatility sometimes jumps to a higher level. Finally, it is of interest to investigate the effects of misspecifying a CLEAR model by a standard linear model.

Appendix

To derive the gradient of the log likelihood (9) function we consider partial derivatives of the unconditional pdf of y_t with respect to the model parameters $\{\mu, \alpha_1, \ldots, \alpha_p, \sigma_\varepsilon, \beta, \sigma_u\}$

$$\frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \mu} = \frac{1}{f(y_t|Y_{t-1}, x_t; \theta)} \Bigg[\Phi\left(\frac{-x_t'\beta}{\sigma_u}\right) \frac{e_t}{\sigma_\varepsilon^2} f(y_t|Y_{t-1}, v_t; \theta)|_{v_t=0} + \int_{-x_t'\beta}^{\infty} \frac{1}{\sigma_u} \phi\left(\frac{u_t}{\sigma_u}\right) \frac{e_t - x_t'\beta - u_t}{\sigma_\varepsilon^2} f(y_t|Y_{t-1}, v_t; \theta)|_{v_t = x_t\beta + u_t} du_t \Bigg],$$

$$\frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \alpha_i} = \frac{1}{f(y_t|Y_{t-1}, x_t; \theta)} \Bigg[\Phi\left(\frac{-x_t'\beta}{\sigma_u}\right) \frac{e_t}{\sigma_\varepsilon^2} f(y_t|Y_{t-1}, v_t; \theta)|_{v_t=0} + \int_{-x_t'\beta}^{\infty} \frac{1}{\sigma_u} \phi\left(\frac{u_t}{\sigma_u}\right) \frac{e_t - x_t'\beta - u_t}{\sigma_\varepsilon^2} f(y_t|Y_{t-1}, v_t; \theta)|_{v_t = x_t\beta + u_t} du_t \Bigg] y_{t-i},$$

for $i = 1, \ldots, p$,

$$\frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \sigma_\varepsilon} = \frac{1}{f(y_t|Y_{t-1}, x_t; \theta)} \Bigg[\Phi\left(\frac{-x_t'\beta}{\sigma_u}\right) \left(\frac{e_t^2}{\sigma_\varepsilon^3} - \frac{1}{\sigma_\varepsilon}\right) f(y_t|Y_{t-1}, v_t; \theta)|_{v_t=0} + \int_{-x_t'\beta}^{\infty} \frac{1}{\sigma_u} \phi\left(\frac{u_t}{\sigma_u}\right) \left(\frac{(e_t - x_t'\beta - u_t)^2}{\sigma_\varepsilon^3} - \frac{1}{\sigma_\varepsilon}\right) f(y_t|Y_{t-1}, v_t; \theta)|_{v_t = x_t\beta + u_t} du_t \Bigg],$$

$$\frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \beta} = \frac{1}{f(y_t|Y_{t-1}, x_t; \theta)} \Bigg[\frac{-1}{\sigma_u} \phi\left(\frac{-x_t'\beta}{\sigma_u}\right) f(y_t|Y_{t-1}, v_t; \theta)|_{v_t=0} + \frac{1}{\sigma_u} \phi\left(\frac{-x_t'\beta}{\sigma_u}\right) f(y_t|Y_{t-1}, v_t; \theta)|_{v_t=0} - \int_{-x_t'\beta}^{\infty} \frac{1}{\sigma_u} \phi\left(\frac{u_t}{\sigma_u}\right) \frac{\varepsilon_t}{\sigma_\varepsilon^2} f(y_t|Y_{t-1}, v_t; \theta)|_{v_t = x_t'\beta + u_t} du_t \Bigg] x_t$$

and

$$\frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \sigma_u} = \frac{1}{f(y_t|Y_{t-1}, x_t; \theta)} \Bigg[\frac{-1}{\sigma_u^2} \phi\left(\frac{-x_t'\beta}{\sigma_u}\right) f(y_t|Y_{t-1}, v_t; \theta)|_{v_t=0} + \int_{-x_t'\beta}^{\infty} \left(\frac{u_t^2}{\sigma_u^4} - \frac{1}{\sigma_u^2}\right) \phi\left(\frac{u_t}{\sigma_u}\right) f(y_t|Y_{t-1}, v_t; \theta)|_{v_t = x_t\beta + u_t} du_t \Bigg],$$

where e_t is defined below (7) and $f(y_t|Y_{t-1}, v_t; \theta)$ and $f(y_t|Y_{t-1}, x_t; \theta)$ are defined in (7) and (8), respectively. The gradient $g_t(y_t|Y_{t-1}, x_t; \theta) = \partial \ln(f(y_t|Y_{t-1}, x_t; \theta))/\partial\theta$ becomes

$$g_t(y_t|Y_{t-1}, x_t; \theta) = \begin{pmatrix} \frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \mu} \\ \frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \alpha_1} \\ \vdots \\ \frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \alpha_p} \\ \frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \sigma_\varepsilon} \\ \frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \beta} \\ \frac{\partial \ln(f(y_t|Y_{t-1}, x_t; \theta))}{\partial \sigma_u} \end{pmatrix},$$

and hence the gradient of the log likelihood function equals

$$G(Y_T|x_T; \theta) = \sum_{t=1}^{T} g_t(y_t|Y_{t-1}, x_t; \theta).$$

To construct standard errors for our maximum likelihood estimator $\hat{\theta}$ we use an estimate of the information matrix H from the scores' average outer product

$$\hat{H} = \frac{1}{T} \sum_{t=1}^{T} g_t(y_t|Y_{t-1}, x_t; \hat{\theta})' g_t(y_t|Y_{t-1}, x_t; \hat{\theta}),$$

see Hamilton (1996, p.132) for a similar approach in Markov switching time series models.

References

Balke, N.S. and T.B. Fomby, 1994, Large Shocks, Small Shocks, and Economic Fluctuations: Outliers in Macroeconomic Time Series, *Journal of Applied Econometrics* **9**, 181–200.

Bianchi, M. and G. Zoega, 1998, Unemployment Persistence: Does the Size of the Shock Matter, *Journal of Applied Econometrics* **13**, 283–304.

Blanchard, O.J. and L. Summers, 1987, Hysteresis in Unemployment, *European Economic Review* **31**, 288–295.

Chen, R., 1995, Threshold Variable Selection in Open-Loop Threshold Autoregressive Models, *Journal of Time Series Analysis* **16**, 461–482.

Clements, M.P. and D. Hendry, 1993, On the Limitations of Comparing Mean Squared Forecast Errors, *Journal of Forecasting* **12**, 617–637.

Diebold, F.X., J. Lee and G.C. Weinbach, 1994, Regime Switching and Endogenous Transition Probablities, in C.P. Hargreaves (editor), *Nonstationary Time Series Analysis and Cointegration*, Oxford University Press, Oxford, chapter 10, pp. 283–302.

Durland, J.M. and T.H. McCurdy, 1994, Duration-Dependent Transitions in a Markov Model, *Journal of Business & Economic Statistics* **12**, 279–288.

Estrella, A. and F.S. Mishkin, 1997, The Predictive Power of the Term Structure of Interest Rates in Europe and the United Staes: Implications for the European Central Bank, *European Economic Review* **41**, 1375–1401.

Estrella, A. and G.A. Hardouvelis, 1991, The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance* **46**, 555–576.

Everitt, B.S. and D.J. Hand, 1981, *Finite Mixture Distributions*, Monographs on Applied Probability and Statistics, Chapman and Hall, London.

Franses, P.H., 1998, Does Seasonality in Unemployment Change with its (Nonlinear) Business Cycle?, *Econometric Institute Report 9809*, Erasmus University Rotterdam.

Granger, C.W.J. and T. Teräsvirta, 1993, *Modelling Nonlinear Economic Relations*, Oxford University Press, Oxford.

Hamilton, J.D., 1989, A New Approach to the Econometric Analysis of Nonstationary Time Series and Business Cycles, *Econometrica* **57**, 357–384.

Hamilton, J.D., 1994, *Time Series Analysis*, Princeton University Press, New Jersey.

Hamilton, J.D., 1996, Specification Testing in Markov-Switching Time-Series Models, *Journal of Econometrics* **70**, 127–158.

Harvey, C.R., 1988, The Real Term Structure and Consumption Growth, *Journal of Financial Economics* **22**, 305–333.

Johnson, N.L. and S. Kotz, 1970, *Distributions in Statistics: Continuous Univariate Distributions*, Houghton Mifflin, Boston.

Skalin, J. and T. Teraesvirta, 1998, Modelling Asymmetries and Moving Equilibria in Unemployment Rates, Working Paper, Stockholm School of Economics.

Teräsvirta, T., 1994, Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models, *Journal of the American Statistical Association* **89**, 208–218.

Teräsvirta, T., 1998, Modeling Economic Relationships with Smooth Transition Autoregressions, in A. Ullah and D.E.A. Giles (editors), *Handbook of Applied Economics*, Marcel Decker, New York.

Tong, H., 1990, *Non-Linear Time Series: A Dynamic System Approach*, Oxford University Press, Oxford.