

The Performance of Alternative Forecasting

Methods for SETAR Models

*Michael P. Clements**

Department of Economics

University of Warwick

Coventry, CV4 7AL

Tel: (01203) 522990

fax: (01203) 523032

E-mail: M.P.Clements@warwick.ac.uk

and

Jeremy Smith

Department of Economics

University of Warwick

Coventry, CV4 7AL

Tel: (01203) 523336

fax: (01203) 523032

E-mail: Jeremy.Smith@warwick.ac.uk

First draft: June 1996

Final version: February 1997

* The first author acknowledges financial support under ESRC grant L116251015. We are grateful to Jan G. de Gooijer and to two anonymous referees whose comments have improved the focus and organisation of the paper.

Abstract

We compare a number of methods that have been proposed in the literature for obtaining h -step ahead minimum mean square error forecasts for SETAR models. These forecasts are compared to those from an AR model. The comparison of forecasting methods is made using Monte Carlo simulation. The Monte Carlo method of calculating SETAR forecasts is generally at least as good as that of the other methods we consider. An exception is when the disturbances in the SETAR model come from a highly asymmetric distribution, when a Bootstrap method is to be preferred.

An empirical application calculates multi-period forecasts from a SETAR model of US GNP using a number of the forecasting methods. We find that whether there are improvements in forecast performance relative to a linear AR model depends on the historical epoch we select, and whether forecasts are evaluated conditional on the regime the process was in at the time the forecast was made.

Keywords: Threshold model, forecasting, simulations.

1. Introduction

The aim of this paper is to compare the forecast performance of a number of methods that can be used to obtain h -step ahead minimum mean square error forecasts for self-exciting threshold autoregressive (SETAR) models. We assess the relative merits of the alternative methods by Monte Carlo simulation, and vary the design of the Monte Carlo to explore the impact of a range of factors (such as non-zero intercepts in the SETAR models, non-Gaussian innovations, etc.).

The motivation for the paper is that the multi-step forecast performance of SETAR models is often not reported in many of the papers that comprise the large and growing body of literature on applying such models to economics data. This is perhaps surprising since one of the main uses of time-series models is for prediction. Moreover, neither in-sample fit, nor the rejection of the null of linearity in a formal test for non-linearity, guarantee that SETAR models (or non-linear models more generally) will forecast more accurately than linear AR models (see, for example, Diebold (1990), and De Gooijer and Kumar (1992), who argue that the evidence for a superior forecast performance of non-linear models is patchy). One factor that has probably retarded the widespread reporting of forecasts from SETAR models is that it is not possible to obtain closed-form analytic expressions for multi-step ahead forecasts and exact numerical solutions require sequences of numerical integrations.

Thus in this paper we evaluate a number of alternative methods that are available that do not involve numerical integration. Their accuracy, ease of use, and robustness to certain non-standard aspects of the forecasting problem, or forms of model-misspecification, are addressed. The five methods we consider are a Monte Carlo simulation method (MC), a Bootstrap method (BS), the normal forecast errors method (NFEM) due to Al-Qassam and

Lane (1989) and adapted to SETAR models by De Gooijer and De Bruin (1997), a multi-step or Dynamic Estimation method (DE), and a ‘naive’ method (SK). Linear AR models are also used as a benchmark for comparison.

De Gooijer and De Bruin (1997) draw together some of the results in the literature on forecasting SETAR processes, and motivate the NFEM from the Chapman-Kolmogorov relations. They also report a comparison of multi-step forecasts and forecast variances produced by the NFEM, the MC method, and an exact method involving numerical integration, for a small number of SETAR models. We extend their study by considering three additional methods and a broader range of SETAR models, which include models with non-Gaussian disturbances. We do not consider numerical integration techniques.

The comparison between the forecasting methods could be made at a number of different knowledge levels concerning the SETAR model. For example, we could assume that: (i) all the parameters of the SETAR model are known; (ii) the autoregressive parameters are unknown, but the number of regimes, thresholds, delay parameters, and autoregressive lag orders are known; (iii) all the parameters are unknown except the lag orders; (iv) all the parameters including the lag orders are unknown. De Gooijer and De Bruin (1997) operate at level (i) – we make the comparison at level (iii), except we assume that the delay lag is also known. We operate at this level following a suggestion of one of the referees. It allows a comparison of *feasible* forecasting methods, but also has a number of other advantages. It enables the SETAR and AR models to be compared. At level (i) the parameters of the AR model would need to be replaced by their pseudo-true values under the SETAR, which may be difficult to calculate. Moreover, it allows the use of the BS method, which is not applicable at level (i). Lai and Zhu (1991) consider multi-step prediction for general non-linear AR models

when the model parameters are unknown, and apply a number of the forecasting methods we consider to the exponential autoregressive model.

The paper is organised as follows. In section 2 we introduce the basic SETAR model and the alternative forecasting methods. Section 3 discusses the Monte Carlo comparison of forecasting methods. Section 4 applies some of the forecasting methods to SETAR models of US GNP. As well as illustrating the performance of the methods, this empirical example is of considerable interest in its own right, following the recent papers of Tiao and Tsay (1994) and Potter (1995), which suggest that US GNP can be usefully modelled as a SETAR process. However, Hansen (1996) finds the evidence for a SETAR structure for US GNP weak. We approach the question from a forecasting perspective, and our findings stress the importance of evaluating forecasts conditional on the state of the economy.

2. The SETAR and AR models

The SETAR model was introduced by Tong (1978), Tong and Lim (1980) and discussed more fully in Tong (1995). In section 2.1 we briefly outline the model, and in section 2.2 we consider methods of obtaining forecasts. Section 2.3 considers the linear AR model.

2.1. A brief description of the SETAR model

It is assumed that a variable, y_t , is a linear autoregression within a regime, but may move between regimes depending on the value taken by a lag of y_t , say y_{t-d} , where d is known as the length of the delay. Formally, y_{t-d} is continuous on the real line, \mathfrak{R} , so that partitioning the real line defines the number of distinct regimes, say q . The process is in the r_h regime when $r_{h-1} \leq y_{t-d} < r_h$, and in that case y_t follows an AR process of order p_h . For the 2-regime case

($q = 2$) the model is denoted by SETAR(2; p_1, p_2), indicating that the process is an AR(p_1) in regime 1 and an AR(p_2) in regime 2:

$$\begin{aligned} y_t &= \alpha_{1,0} + \sum_{j=1}^{p_1} \alpha_{1,j} y_{t-j} + \varepsilon_{1,t}, & y_{t-d} \leq r \\ y_t &= \alpha_{2,0} + \sum_{j=1}^{p_2} \alpha_{2,j} y_{t-j} + \varepsilon_{2,t}, & y_{t-d} > r \end{aligned} \quad (1a)$$

where $t = 1, \dots, n$ and $\varepsilon_{h,t} \sim G(0, \sigma_h^2)$, $h = 1, 2$. The distribution of the disturbances, $G(\cdot)$, may be Gaussian, but this is not necessarily the case, and in section 3 we consider a number of possibilities. The model can be written more compactly using the indicator function as:

$$y_t = \alpha_0 + \sum_{j=1}^{p_1} \alpha_j y_{t-j} + \varepsilon_{1,t} + I_{t-d}(r) (\beta_0 + \sum_{j=1}^{p_2} \beta_j y_{t-j} + \varepsilon_{2,t} - \varepsilon_{1,t}) \quad (1b)$$

where, $I_t(r) = I(y_t > r) = \begin{cases} 1 & \text{if } y_t > r \\ 0 & \text{otherwise} \end{cases}$, so that comparing equations (1a) and (1b):

$\alpha_{1,i} = \alpha_i$, $i = 1, \dots, p_1$, and $\alpha_{2,i} = \alpha_i + \beta_i$, $i = 1, \dots, p_2$. When $p_1 \neq p_2$, some parameters will be zero.

The necessary and sufficient conditions for stationarity of SETAR models only exist in the literature for a number of special cases (see, for example, De Gooijer and De Bruin, 1997).

2.2. Methods for forecasting SETAR models

In this section, we briefly outline five forecasting methods that will be investigated in section 3. All the methods are discussed in Granger and Teräsvirta (1993) in the context of forecasting from a general non-linear model. For expositional simplicity, the discussion assumes a SETAR(2;1,1) model and that the delay parameter is one ($d = 1$).

The normal forecast error method (NFEM):

The first forecasting method we consider is the normal forecast errors method (NFEM) suggested by Al-Qassam and Lane (1989) for a first-order exponential autoregressive model and extended by De Gooijer and De Bruin (1997) to SETAR(q; p₁, p₂) with general d. This method essentially takes a weighted average of the forecasts from regime 1,

$\hat{y}_{1,n+k} = \alpha_{1,0} + \alpha_{1,1}\hat{y}_{n+k-1}^M$ and regime 2, $\hat{y}_{2,n+k} = \alpha_{2,0} + \alpha_{2,1}\hat{y}_{n+k-1}^M$, to form the forecast as:

$$\hat{y}_{n+k}^M = p_{k-1}\hat{y}_{1,n+k} + (1-p_{k-1})\hat{y}_{2,n+k} + (\alpha_{2,1} - \alpha_{1,1})\hat{\sigma}_{n+k-1}\phi\left(\frac{r - \hat{y}_{n+k-1}^M}{\hat{\sigma}_{n+k-1}}\right) \quad k = 2, \dots \quad (2)$$

where the weight p_{k-1} is the probability of the process being in the lower regime at time $n+k-1$ assuming normality, that is $p_{k-1} = \Phi\{(r - \hat{y}_{n+k-1}^M) / \hat{\sigma}_{n+k-1}\}$. Substituting for $\hat{y}_{1,n+k}$, $\hat{y}_{2,n+k}$ and p_{k-1} in (2) we obtain the following recursive relation for obtaining approximate k -step ($k > 1$) forecasts:

$$\begin{aligned} \hat{y}_{n+k}^M &= \Phi\left(\frac{r - \hat{y}_{n+k-1}^M}{\hat{\sigma}_{n+k-1}}\right)(\alpha_{1,0} + \alpha_{1,1}\hat{y}_{n+k-1}^M) + \Phi\left(-\frac{r - \hat{y}_{n+k-1}^M}{\hat{\sigma}_{n+k-1}}\right)(\alpha_{2,0} + \alpha_{2,1}\hat{y}_{n+k-1}^M) \\ &+ \phi\left(\frac{r - \hat{y}_{n+k-1}^M}{\hat{\sigma}_{n+k-1}}\right)(\alpha_{2,1} - \alpha_{1,1})\hat{\sigma}_{n+k-1} \quad k = 2, \dots \end{aligned} \quad (3)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the distribution and probability density function for a standard $N(0,1)$. The recursion begins with:

$$\hat{y}_{n+1}^M = \alpha_0 + \alpha_1 y_n + I_n(r)(\beta_0 + \beta_1 y_n).$$

Equation (3) makes use of $\hat{\sigma}_{n+k-1}$, the $(n+k-1)^{\text{th}}$ -step ahead forecast standard error. This is calculated from the recursive formula:

$$\begin{aligned} \hat{\sigma}_{n+k}^2 = & \{(\alpha_{1,0} + \alpha_{1,1}\hat{y}_{n+k-1}^M)^2 + \alpha_{1,1}^2 \hat{\sigma}_{n+k-1}^2\}p_{k-1} + \{(\alpha_{2,0} + \alpha_{2,1}\hat{y}_{n+k-1}^M)^2 + \alpha_{2,1}^2 \hat{\sigma}_{n+k-1}^2\}(1-p_{k-1}) \\ & + \left\{ \alpha_{2,1}^2 (r - \hat{y}_{n+k-1}^M) + 2\alpha_{2,1}(\alpha_{2,0} + \alpha_{2,1}\hat{y}_{n+k-1}^M) - (\alpha_{1,1}^2 (r - \hat{y}_{n+k-1}^M) + 2\alpha_{1,1}(\alpha_{1,0} + \alpha_{1,1}\hat{y}_{n+k-1}^M)) \right\} \hat{\sigma}_{n+k-1} p_{k-1} \\ & + \sigma_\varepsilon^2 - (\hat{y}_{n+k}^M)^2 \end{aligned}$$

which is applicable for a model with a regime-invariant variance of the disturbance term, σ_ε^2 .

De Gooijer and De Bruin (1997) show how the method can be extended to models with $q > 2$, $p_1 = p_2 > 1$ and $d > 1$. The NFEM may not be appropriate when the underlying disturbances are not normal – see, for example, Lai and Zhu (1991), p.318.

The Monte Carlo method (MC):

The Monte Carlo method (denoted MC) is a simple simulation method for obtaining multi-step forecasts that can be applied as easily to complex models (high autoregressive lag orders, several regimes) as to the simple SETAR(2;1,1) model. For forecasting the $(n+1)^{\text{th}}$ observation conditional on y_n , the regime is known with certainty, so that the 1-step MC forecast is simply:

$$\hat{y}_{n+1}^{\text{MC}} = \alpha_0 + \alpha_1 y_n + I_n(r)(\beta_0 + \beta_1 y_n),$$

which is identical to the 1-step forecast produced by the NFEM. However, for observation $n+2$ the forecast regime is determined by y_{n+1} , but we only have $\hat{y}_{n+1}^{\text{MC}}$ (which differs from y_{n+1} by an error term). Define the forecasts for y_{n+2} , y_{n+3} and y_{n+k} for a realisation of the error process, $\zeta_{i,j}^h$, as:

$$\hat{y}_{n+2}^{\text{MCj}} = \alpha_0 + \alpha_1 \hat{y}_{n+1}^{\text{MC}} + I_{n+1}(r)\{\beta_0 + \beta_1 \hat{y}_{n+1}^{\text{MC}}\} + \zeta_{2,j}^h \quad (4)$$

$$\hat{y}_{n+3}^{\text{MCj}} = \alpha_0 + \alpha_1 \hat{y}_{n+2}^{\text{MCj}} + I_{n+2}(r)\{\beta_0 + \beta_1 \hat{y}_{n+2}^{\text{MCj}}\} + \zeta_{3,j}^h \quad (5)$$

and:

$$\hat{y}_{n+k}^{\text{MCj}} = \alpha_0 + \alpha_1 \hat{y}_{n+k-1}^{\text{MCj}} + I_{n+k-1}(r)\{\beta_0 + \beta_1 \hat{y}_{n+k-1}^{\text{MCj}}\} + \zeta_{k,j}^h \quad (6)$$

where, $I_{n+k-1}(r) = I(\hat{y}_{n+k-1}^{MCj} > r) = \begin{cases} 1 & \text{if } \hat{y}_{n+k-1}^{MCj} > r \\ 0 & \text{otherwise} \end{cases}$ and $k > 1$. Now, averaging these forecasts

across the $j = 1, \dots, N$ iterations of the Monte Carlo yields:

$$\hat{y}_{n+k}^{MC} = \frac{1}{N} \sum_{j=1}^N \hat{y}_{n+k}^{MCj}, \quad (7)$$

for $k > 1$. Notice that the pseudo-random numbers are super-scripted by the regime, so that the drawing in period $n+k$ is made from a distribution with a variance appropriate for the regime the process is in (determined by the forecast value of the process for $n+k-1$). We assume that the $\zeta_{i,j}^h$ are independent Gaussian variates, $\zeta_{i,j}^h \sim N(0, \sigma_h^2)$. The normality assumption is used here regardless of the actual distribution of $\varepsilon_{h,t}$. In the Monte Carlo study of section 3 we assess the performance of the MC method for some non-Gaussian distributions of $\varepsilon_{h,t}$. This method is used extensively by Clements and Smith (1996) to obtain forecasts for a number of empirical SETAR models.

The Skeleton method (SK):

The third method can be viewed as a special case of the MC method, where the $\zeta_{i,j}^h$ are set to zero for all i, h . This ‘naive’ or ‘skeleton’ method (SK) (see Tong, 1995) amounts to approximating the expectation of a function of random variables by that function of the expectation.

The Bootstrap method (BS):

The fourth method is the bootstrap method (BS), which is naturally closely related to the MC method. The forecasts are constructed in similar manner to the MC forecasts except that the

$\zeta_{i,j}^h$ are drawn randomly (with replacement) from the within-sample, regime-specific residuals, rather than the normal distribution that is used in the MC method.

The Dynamic Estimation method (DE):

The fifth method is motivated by the dynamic estimation (DE) method discussed by Granger and Teräsvirta (1993). Consider the usual method of estimation of the model given by equation (1). Conditional on the regime threshold r and the delay lag d being known, the sample can simply be split into two groups and an OLS regression can be run on the observations belonging to each regime separately. Alternatively, indicator functions can be used in a single regression, constraining the residual error variance to be constant across regimes (see, for example, Potter, 1995, p.113.). When r and d are unknown, the model is estimated by searching over these two parameters – that is, the model is estimated by OLS at each pair of values (r,d) . The range of values of r is usually restricted to those between the 15th and 85th percentile of the empirical distribution of y_{t-d} , following Andrews (1993) and Hansen (1996), while d typically takes on the values $0,1,2,\dots$ up to the maximum lag length allowed. For a known lag order, the selected model is that for which the pair (r,d) minimize the overall residual sum of squares (RSS, equal to the sum of the RSS in each regime). This is the model estimation method we adopt in section 3 for all the forecasting methods other than DE, except to simplify matters we set d equal to unity. It is a ‘1-step’ estimation method, since OLS minimizes the sum of squares of 1-step errors.

In the spirit of dynamic estimation applied to linear models (see, for example, Clements and Hendry, 1996, for references), for forecasting k periods ahead, we forecast y_{n+k} directly from information known at period n (and earlier) by using the in-sample relationship between y_t and y_{t-k} . For example, for a 2-step ahead forecast horizon, the model is estimated by minimizing the

in-sample sum of squares of 2-step errors. From (1b) the 2-step version of the model can be written as:

$$y_{t+2} = \alpha_0 + \alpha_1(\alpha_0 + \alpha_1 y_t + I_t(r)(\beta_0 + \beta_1 y_t) + \varepsilon_{t+1}) + I_{t+1}(r)\{\beta_0 + \beta_1(\alpha_0 + \alpha_1 y_t + I_t(r)(\beta_0 + \beta_1 y_t) + \varepsilon_{t+1})\} + \varepsilon_{t+2}. \quad (8)$$

If we ignore terms multiplied by $I_{t+1}(r)$ in equation (8) and simply relate y_{t+2} to a two-regime model based on y_t we obtain:

$$y_{t+2} = \tilde{\alpha}_{0,2} + \tilde{\alpha}_{1,2} y_t + I_t(\tilde{r}_2)(\tilde{\beta}_{0,2} + \tilde{\beta}_{1,2} y_t) + \tilde{\varepsilon}_{t+2} \quad (9)$$

the parameters of which can be estimated by minimising the sum of squares of $\tilde{\varepsilon}_{t+2}$ over $\tilde{\alpha}_{i,2}$, $\tilde{\beta}_{i,2}$, $i = 0, 1$, and \tilde{r}_2 , for the sample period $t = 1, \dots, n$. This requires a ‘grid-search’ over \tilde{r}_2 as for ‘1-step’ estimation. The generalisation to k -step forecasts entails minimising the sum of squares of $\tilde{\varepsilon}_{t+k}$ over $\tilde{\alpha}_{i,2}$, $\tilde{\beta}_{i,2}$, $i = 0, 1$, and \tilde{r}_k , where:

$$y_{t+k} = \tilde{\alpha}_{0,k} + \tilde{\alpha}_{1,k} y_t + I_t(\tilde{r}_k)(\tilde{\beta}_{0,k} + \tilde{\beta}_{1,k} y_t) + \tilde{\varepsilon}_{t+k} \quad (10)$$

The forecast function for y_{n+k} is linear in the estimated parameters and y_n (and $I_n(\hat{r}_k)y_n$):

$$\hat{y}_{n+k}^{DE} = \hat{\alpha}_{0,k} + \hat{\alpha}_{1,k} y_n + I_n(\hat{r}_k)(\hat{\beta}_{0,k} + \hat{\beta}_{1,k} y_n) \quad (11)$$

where $\hat{\alpha}_{i,k}$, $\hat{\beta}_{i,k}$ and \hat{r}_k are the least squares estimates.

In a linear setting, dynamic estimation would not be expected to yield gains in terms of forecast accuracy over the traditional minimization of 1-step in-sample errors (OLS) when the model is correctly-specified. Here the model is correctly-specified for the data generating

process, but gains might be expected to accrue from the simpler means of obtaining forecasts from models estimated by DE, and also when the disturbances are non-Gaussian.

An obvious drawback with this method is the computational cost of estimating the model for each forecast horizon.

2.3. The linear AR model

Finally, we consider forecasting from linear AR models, which we denote by AR(p), where p is the lag length. Fitting a linear AR(p) process to y_t yields a linear predictor, where the forecasts solve the recursive formula:

$$\hat{y}_{n+k}^{\text{AR}(p)} = \hat{\phi}_0 + \sum_{j=1}^p \hat{\phi}_j \hat{y}_{n+k-j}^{\text{AR}(p)}, \quad (12)$$

with $\hat{y}_{n+k}^{\text{AR}(p)} = y_{n+k}$ for $k \leq 0$, and the $\hat{\phi}_j$ are the coefficient estimates from fitting an AR(p) (with an intercept) to the series y_t , $t = 1, \dots, n$.

3. Monte Carlo design

The forecast method comparisons reported in this paper are for the SETAR(2;1,1) model with $d = 1$:

$$\begin{aligned} y_t &= \alpha_{1,0} + \alpha_{1,1}y_{t-1} + \varepsilon_{1,t} & y_{t-1} \leq r \\ y_t &= \alpha_{2,0} + \alpha_{2,1}y_{t-1} + \varepsilon_{2,t}, & y_{t-1} > r \end{aligned} \quad (13)$$

where $\varepsilon_{h,t} \sim G(0, \sigma_h^2)$. We consider (13) for various combinations of the parameters drawn

from the parameter spaces: $(\alpha_{1,1}, \alpha_{2,1}) \in (-0.8, -0.6, -0.4, -0.2, 0.2, 0.6, 0.8)$ and

$(\alpha_{1,0}, \alpha_{2,0}) = (-0.5, -0.25, 0.0, 0.25, 0.5)$, with G, the distribution of the error term, being either

normal (N), uniform (U) or chi-squared(1), ($\chi^2(2)$). For the results reported in this section, d is assumed to be known to be unity.

The comparison of the various forecasting methods for the SETAR model is made by Monte Carlo as follows. For each iteration of the Monte Carlo we simulate a series of 210 observations from (13). The first 200 observations are used to estimate the SETAR and AR models ($n = 200$). The lag orders p_1 and p_2 are assumed to be known to be 1, and the parameters $\{\alpha_{1,0}, \alpha_{1,1}, \alpha_{2,0}, \alpha_{2,1}, \sigma_h^2, r\}$ are all estimated including the analogues for DE. We impose the stationarity conditions for a SETAR(2;1,1) model, that is: $\alpha_{1,1} < 1$, $\alpha_{2,1} < 1$, and $\alpha_{1,1}\alpha_{2,1} < 1$ (and similarly for the DE estimates). Since forecast comparison is by MSFE, failure to do so might result in the estimated SETAR models generating explosive forecasts on a small number of Monte Carlo replications which would inflate the calculated MSFEs and make interpretation of the results more difficult.

1 to 10-step ahead forecasts of the 10 remaining observations are then generated from the estimated SETAR model for each of the methods described in section 2, and from the estimated AR models. For the MC method this entails a set of Monte Carlo simulations consisting of $N=500$ replications for each iteration. The forecast errors are then calculated and stored for each forecasting method and model. The above is then repeated for each of 1000 iterations to generate a Monte Carlo sample of forecast errors, and the Monte Carlo estimates of the mean squared forecast errors (MSFEs) are the sums of the squares of the forecast errors over the 1000 iterations. Notice that the model is re-estimated on each iteration. Moreover, for the DE method the model is estimated for each forecast horizon, on each replication. Only first and second-order AR models are considered, since preliminary investigations suggested that longer lag orders do not yield an improved forecast performance.

The results are organised to highlight the impact of some of these aspects of the design of the experiments on the outcomes of the forecast comparisons. We consider the relative performance of the methods for the SETAR model, and also draw out the gains (if any) to using the SETAR over an AR for forecasting. We do not consider the impact of changing the variance of the SETAR model disturbances – De Gooijer and De Bruin (1997) document the deterioration in the NFEM as this variance increases.

3.1. Results

3.1.1. Gaussian disturbances and zero intercepts

The first set of experiments focus on the SETAR model slope parameters $\alpha_{1,1}, \alpha_{2,1}$. We assume that the disturbances are Gaussian, $\varepsilon_{h,t} \sim N(0, \sigma_h^2)$, with $\sigma_1 = \sigma_2 = 0.5$, and set the intercepts to zero, $\alpha_{1,0} = \alpha_{2,0} = 0$. A selection of our results for the MSFEs of the AR(1), NFEM, BS, SK and DE, all divided by MC, are reported in Table 1 for 1, 2 and 5-step ahead forecasts. The results for the AR(2) were never better than for the AR(1) and hence are not reported. It is apparent that there is little to choose between the MC, NFEM and BS methods for this experimental design. However, the DE method is worse than the BS, MC and NFEM for this set of experiments, and using the ‘naive method’ SK yields much worse forecasts in some instances. There appears to be no tendency for the SK MSFEs to converge on those of the other methods as the horizon lengthens, chiefly due to the bias.

Table 1

MSFEs of SETAR forecast methods and AR(1) model relative to MC method

($\alpha_{1,0} = \alpha_{2,0} = 0$, $\sigma_1 = \sigma_2 = 0.5$, Gaussian disturbances)

Design	Forecast horizon	AR(1)	NFEM	BS	SK	DE
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = -0.6$	1	0.99	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.01	1.03	1.03
	5	1.00	1.00	1.01	1.05	1.04
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = 0.6$	1	1.09	1.00	1.00	1.00	1.00
	2	1.02	1.01	1.01	1.10	1.03
	5	1.00	1.00	1.00	1.19	1.05
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = -0.4$	1	0.98	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.01	1.04	1.02
	5	1.00	1.00	1.00	1.04	1.05
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = 0.8$	1	1.09	1.00	1.00	1.00	1.00
	2	1.04	1.00	1.00	1.08	1.03
	5	1.01	1.00	1.00	1.16	1.05
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = -0.6$	1	0.98	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.04	1.04
	5	1.00	1.00	1.01	1.03	1.03
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = 0.6$	1	1.03	1.00	1.01	1.00	1.00
	2	1.01	1.00	1.00	1.06	1.02
	5	1.00	1.00	1.00	1.13	1.04
$\alpha_{1,1} = -0.2$ $\alpha_{2,1} = 0.2$	1	0.97	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.01	1.04	1.03
	5	1.00	1.00	1.01	1.02	1.04
$\alpha_{1,1} = 0.2$ $\alpha_{2,1} = -0.2$	1	0.98	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.01	1.05	1.03
	5	1.00	1.00	1.00	1.04	1.05
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.8$	1	1.18	1.00	1.00	1.00	1.00
	2	1.02	1.00	1.01	1.10	1.05
	5	1.00	1.00	1.00	1.21	1.02
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.4$	1	1.10	1.00	1.00	1.00	1.00
	2	1.02	1.00	1.00	1.09	1.05
	5	1.00	1.00	1.00	1.13	1.05
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = 0.4$	1	0.98	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.03	1.05
	5	1.00	1.00	1.00	1.07	1.04

The gains relative to an AR model can be large when $\alpha_{1,1}$ and $\alpha_{2,1}$ are of opposite sign, and close to the unit circle. Not surprisingly, the AR model is penalised less heavily when these parameters are of the same sign and close together, since the process generating the data is then closer to a single-regime (i.e. linear) model. The closer these parameters are to zero, the closer the process to white noise, which can be optimally forecast by an AR model with AR coefficients of zero.

More interestingly, the penalty to using an AR model is greater when $\alpha_{1,1} > 0$ and $\alpha_{2,1} < 0$ than when these parameters have the reverse signs (compare, for example, the entry in Table 1 for $\alpha_{1,1} = 0.6$ and $\alpha_{2,1} = -0.8$ with that for $\alpha_{1,1} = -0.6$ and $\alpha_{2,1} = 0.8$).

When $\alpha_{1,1} < 0$ and $\alpha_{2,1} > 0$ (e.g., $\alpha_{1,1} = -0.6$ and $\alpha_{2,1} = 0.8$), once the process enters the lower regime due to a sufficiently large negative shock $\varepsilon_t \leq -\alpha_{2,1}y_{t-1}$, the process will flip back to the upper regime the next period unless $\varepsilon_{t+1} \leq -\alpha_{1,1}y_t$. This ‘anti-persistent’ behaviour can be exploited by the SETAR model to yield improved forecasts. It is similar when $\alpha_{1,1} > 0$ and $\alpha_{2,1} < 0$, in which case the process will tend not to remain in the upper regime, which also suggests improved forecasts from the SETAR model.

Table 1 indicates that any gains to the SETAR relative to the AR model will be negligible after 5-steps ahead, and we can do no better than forecasting the unconditional (zero) mean of the process.

3.1.2. Uniform disturbances

We investigated the impact of uniform disturbances in (13), $\varepsilon_h / \sigma_h \sim \{U[0,1] - 0.5\}\sqrt{12}$, for a range of values of α_1^1 and α_1^2 (with zero intercepts, $\alpha_{1,0} = \alpha_{2,0} = 0$). The NFEM is only

strictly applicable when the disturbances of the SETAR process are normally distributed, and De Gooijer and De Bruin (1997) conjecture that it may be possible to adapt the method for other distributions. However, it is of interest to ascertain the robustness of the NFEM to various departures from normality, since in practical applications this assumption will typically only hold approximately.

Table 2 reports a selection of results for the same values of $\alpha_{1,1}$ and $\alpha_{2,1}$ as in Table 1 for the Gaussian case. It is evident from Table 2 that there are no gains to the BS method relative to the MC method. Although not shown, the NFEM performs badly at the longer horizons (8 to 10-steps ahead) for experiments where one or both of $\alpha_{1,1}$ and $\alpha_{2,1}$ are close to zero. The problem appears to be that the recursive formula for the forecast error variance starts to generate explosive values.

Table 2

MSFEs of SETAR forecast methods and AR(1) model relative to MC method

($\alpha_{1,0} = \alpha_{2,0} = 0$, $\sigma_1 = \sigma_2 = 0.5$, Uniform disturbances)

Design	Forecast horizon	AR(1)	NFEM	BS	SK	DE
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = -0.6$	1	0.98	1.00	1.00	1.00	1.00
	2	0.99	0.99	0.99	1.01	1.04
	5	1.00	1.00	1.00	1.01	1.02
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = 0.6$	1	1.08	1.00	1.01	1.00	1.00
	2	1.00	1.00	1.00	1.09	1.04
	5	1.00	1.00	1.00	1.20	1.03
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = -0.4$	1	0.96	1.00	1.00	1.00	1.00
	2	0.99	0.99	1.00	1.02	1.05
	5	1.00	0.99	0.99	1.02	1.03
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = 0.8$	1	1.07	1.00	1.00	1.00	1.00
	2	1.03	1.00	1.00	1.08	1.01
	5	1.00	1.01	1.01	1.23	1.05
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = -0.6$	1	0.98	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.02	1.05
	5	1.00	1.00	1.00	1.02	1.02
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = 0.6$	1	1.03	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.07	1.01
	5	1.00	1.00	1.00	1.14	1.03
$\alpha_{1,1} = -0.2$ $\alpha_{2,1} = 0.2$	1	0.99	1.00	1.00	1.00	1.00
	2	0.99	1.00	1.00	1.02	1.04
	5	1.00	1.00	1.00	1.06	1.03
$\alpha_{1,1} = 0.2$ $\alpha_{2,1} = -0.2$	1	0.97	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.00	1.06	1.03
	5	1.00	1.00	1.00	1.04	1.02
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.8$	1	1.06	1.00	1.00	1.00	1.00
	2	0.99	1.00	0.99	1.14	1.04
	5	1.00	1.00	1.00	1.24	1.04
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.4$	1	1.00	1.00	1.00	1.00	1.00
	2	0.99	1.00	0.99	1.09	1.03
	5	1.00	1.00	1.00	1.15	1.02
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = 0.4$	1	0.97	1.00	1.00	1.00	1.00
	2	0.99	1.00	1.00	1.03	1.00
	5	1.00	1.00	1.01	1.09	1.03

Relative to the AR(1) model, the large gains to the SETAR in the Gaussian case, when $\alpha_{1,1}$ and $\alpha_{2,1}$ are large and $\alpha_{1,1} > 0$ and $\alpha_{2,1} < 0$, are reduced, but the (smaller) gains to the SETAR when the slope parameters take on the reverse signs are affected less by the move from Gaussian to uniform disturbances.

3.1.3. Chi-squared disturbances

We also study the impact of asymmetric disturbances in the SETAR model, with the errors in equation (13) drawn from a χ^2 -distribution, $\varepsilon_h / \sigma_h \sim \{\chi^2(2) - 2\} / \sqrt{4}$. Table 3 reports a selection of results for the same values of $\alpha_{1,1}$ and $\alpha_{2,1}$ as in Tables 1 and 2 for ease of comparison (and again zero intercepts). It is readily apparent that in some cases there are gains to BS relative to MC, but these are rarely much larger than 2%, and occur for values of $\alpha_{1,1}$ and $\alpha_{2,1}$ closer to zero. In fact the gain to using the BS method is maximized at $\alpha_{1,1} = \alpha_{2,1} = 0$ (not shown), where the process is a series of iid χ^2 variates. The NFEM sometimes produces very large numbers at the longer horizons (not shown) due to the forecast variance exploding, occasioned by $\hat{\alpha}_{i,1} < -1$, $i = 1, 2$.

Table 3

MSFEs of SETAR forecast methods and AR(1) model relative to MC method

($\alpha_{1,0} = \alpha_{2,0} = 0$, $\sigma_1 = \sigma_2 = 0.5$, $\chi^2(2)$ disturbances)

Design	Forecast horizon	AR(1)	NFEM	BS	SK	DE
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = -0.6$	1	0.97	1.00	1.00	1.00	1.00
	2	0.986	1.00	1.00	1.02	1.01
	5	0.99	1.00	1.01	1.04	1.02
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = 0.6$	1	1.07	1.00	1.00	1.00	1.00
	2	1.00	1.00	0.98	1.05	1.02
	5	0.99	-	0.99	1.11	1.00
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = -0.4$	1	0.95	1.00	1.00	1.00	1.00
	2	0.98	1.00	1.00	1.01	1.01
	5	1.00	1.00	1.00	1.03	1.02
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = 0.8$	1	1.03	1.00	1.01	1.00	1.00
	2	1.02	1.00	0.98	1.02	1.00
	5	0.98	-	0.98	1.12	1.02
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = -0.6$	1	0.96	1.00	1.01	1.00	1.00
	2	0.99	1.00	1.01	1.03	1.01
	5	1.00	1.00	1.01	1.02	1.02
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = 0.6$	1	1.03	1.00	1.00	1.00	1.00
	2	0.99	1.00	0.98	1.03	1.00
	5	0.99	-	0.98	1.09	1.01
$\alpha_{1,1} = -0.2$ $\alpha_{2,1} = 0.2$	1	0.98	1.00	1.00	1.00	1.00
	2	0.95	0.99	0.98	0.96	0.99
	5	1.00	-	1.00	1.03	1.03
$\alpha_{1,1} = 0.2$ $\alpha_{2,1} = -0.2$	1	1.00	1.00	1.00	1.00	1.00
	2	0.98	1.00	0.99	1.04	1.01
	5	0.98	-	0.99	1.00	1.02
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.8$	1	1.24	1.00	1.00	1.00	1.00
	2	1.02	1.00	1.00	1.10	1.06
	5	1.00	1.00	1.00	1.21	1.02
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.4$	1	1.08	1.00	1.00	1.00	1.00
	2	1.01	1.00	1.00	1.09	1.05
	5	0.99	1.00	1.00	1.11	1.03
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = 0.4$	1	0.97	1.00	1.00	1.00	1.00
	2	1.00	1.00	1.01	1.03	1.02
	5	1.00	1.00	1.00	1.05	1.03

The AR model dominates the SETAR when the slope parameters are of the same sign (e.g.

$\alpha_{1,1} = -0.6, \alpha_{2,1} = -0.4$) or are of opposite sign but close to zero (e.g.

$\alpha_{1,1} = -0.2, \alpha_{2,1} = 0.2$).

3.1.4. Non-zero intercepts

The effect of allowing non-zero intercepts in the SETAR model (13) was investigated for the case of Gaussian disturbances. For most of the combinations of intercepts ($\alpha_{1,0}, \alpha_{2,0}$) and slope parameters ($\alpha_{1,1}, \alpha_{2,1}$) we considered, the MC, BS and NFEM all performed almost equally as well. For the larger values of the intercepts (± 0.5), the NFEM was sometimes a little worse than MC and BS at longer horizons, but any such differences were usually only of the order of 1 or 2%. The entry in Table 4 is omitted for the experiment with $\alpha_0^1 = \alpha_0^2 = 0.5$ and $\alpha_1^2 = 0.8$ because there were insufficient observations in the lower regime to reliably estimate the SETAR model. In Table 4 we report the results for the AR(1) relative to the MC method, and focus the discussion on the comparison of the SETAR and AR models.

Table 4
MSFE for AR(1) relative to MC method
($\sigma_1 = \sigma_2 = 0.5$, Gaussian disturbances)

Design	Forecast horizon	$\alpha_{1,0}, \alpha_{2,0}$					
		.25, .25	-.25, .25	.25, -.25	.5, .5	-.5, .5	.5, -.5
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = -0.6$	1	0.99	1.05	1.06	0.98	1.30	1.14
	2	1.00	1.00	1.04	1.00	1.00	1.23
	5	1.00	1.00	1.05	1.01	1.00	1.37
$\alpha_{1,1} = -0.8$ $\alpha_{2,1} = 0.6$	1	1.02	1.02	1.41	0.95	1.00	1.87
	2	1.01	0.99	1.02	1.00	1.00	1.13
	5	1.00	1.00	1.00	0.99	1.00	1.00
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = -0.4$	1	0.98	1.06	1.05	0.97	1.32	1.18
	2	1.00	1.01	1.02	1.00	1.02	1.20
	5	1.00	1.00	1.01	1.00	1.00	1.15
$\alpha_{1,1} = -0.6$ $\alpha_{2,1} = 0.8$	1	0.98	1.04	1.41	-	0.98	1.71
	2	1.00	1.00	1.02	-	0.99	1.07
	5	1.00	1.00	1.00	-	1.01	1.00
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = -0.6$	1	0.98	1.04	1.06	0.99	1.31	1.17
	2	1.00	1.00	1.04	1.00	1.02	1.16
	5	1.00	1.00	1.01	1.00	1.00	1.10
$\alpha_{1,1} = -0.4$ $\alpha_{2,1} = 0.6$	1	1.00	1.02	1.24	0.98	0.97	1.63
	2	1.00	0.99	1.00	0.99	0.98	1.11
	5	1.00	1.00	1.00	1.00	1.02	1.00
$\alpha_{1,1} = -0.2$ $\alpha_{2,1} = 0.2$	1	0.97	1.08	1.11	0.96	1.23	1.44
	2	1.00	1.01	1.02	1.00	1.09	1.12
	5	1.00	1.00	1.00	1.00	1.01	1.01
$\alpha_{1,1} = 0.2$ $\alpha_{2,1} = -0.2$	1	1.00	1.075	1.10	0.98	1.19	1.30
	2	1.00	1.00	1.02	1.00	1.05	1.15
	5	1.00	1.00	1.00	1.00	1.01	1.01
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.8$	1	1.24	1.07	1.40	1.20	0.97	1.86
	2	1.00	1.02	1.01	1.00	0.99	1.19
	5	1.00	1.00	1.00	1.00	0.99	1.00
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = -0.4$	1	1.11	1.07	1.21	1.08	0.99	1.57
	2	1.00	1.01	0.99	1.00	0.99	1.14
	5	1.00	1.00	1.01	1.00	1.00	1.00
$\alpha_{1,1} = 0.6$ $\alpha_{2,1} = 0.4$	1	0.98	1.02	1.10	0.96	1.10	1.30
	2	1.00	1.02	1.00	0.99	1.16	1.01
	5	1.00	1.00	1.01	1.00	1.11	1.00

Non-zero intercepts can greatly extenuate or attenuate the relative forecast performance of the SETAR model. The pattern of results in Table 4 can be understood by running through a number of cases. Consider the case of $\alpha_{1,1} = -0.8$ and $\alpha_{2,1} = 0.6$, which yielded a 1-step gain to the SETAR of 9% in the absence of intercepts (see Table 1). When $\alpha_{1,0} = \alpha_{2,0} = 0.25$ the process will be almost exclusively in the upper regime, (and will only move to the lower regime if $\varepsilon_t \leq -(0.6y_{t-1} + 0.25)$) and the gains to the SETAR model are consequently reduced (and become negative when the intercepts are both equal to 0.5). If instead $-\alpha_{1,0} = \alpha_{2,0} = 0.25$ the previous argument is little changed. When, however, $\alpha_{1,0} = -\alpha_{2,0} = 0.25$, the process will tend to oscillate between regimes, exhibiting a pattern that can be exploited by the SETAR to record large gains (40%) relative to the AR model. This pattern of movement between regimes is heightened when $\alpha_{1,0} = -\alpha_{2,0} = 0.5$, and the gain at 1-step ahead increases to 85%.

As another example, consider the case of $\alpha_{1,1} = 0.6$ and $\alpha_{2,1} = -0.8$. When the intercepts are zero there are gains to the SETAR of nearly 18% at 1-step ahead (see Table 1). Unlike in the case just considered (when the signs of the AR parameters are reversed), setting $\alpha_{1,0} = \alpha_{2,0} = 0.25$, (or 0.5), does not diminish the gains to the SETAR model, since $\alpha_{2,1} < 0$ ensures the process will still be observed in the lower regime. As for the previous case, the gains increase when $\alpha_{1,0} = -\alpha_{2,0} = 0.25$, or 0.5, when the process again oscillates between the regimes. When $\alpha_{1,0} = -\alpha_{2,0} = -0.5$ the gains to the SETAR become negative because the process remains in the lower regime most of the time.

There can be substantial gains to the SETAR even when $\alpha_{1,1} = \alpha_{2,1}$ provided the intercepts are different and the process oscillates between regimes.

4. SETAR and AR model forecasts of post-war US GNP

In this section some of the forecasting methods are used to generate multi-step forecasts from SETAR models of US GNP. Potter (1995) fitted a SETAR(2;5,5) model to seasonally-adjusted real US GNP data for the period 1947:1-1990:4, with the coefficients on the third and fourth-order lags set to zero. Tiao and Tsay (1994) estimated a two-regime model as well as a more elaborate four-regime model, and obtained qualitatively similar results to Potter (1995). These models suggest that US GNP is relatively robust to negative shocks – the economy moves quickly out of recession. Tiao and Tsay (1994) find that their SETAR model forecasts (calculated by the MC method) are only a little better than those from an AR(2) model, but are markedly superior if we only consider forecasts which are made when the economy is in the lower regime. This reflects the ability of the SETAR model to capture the rapid movement out of recession.

In section 4.1 we redo the Tiao and Tsay forecast exercise with some minor but nevertheless important modifications to evaluate the performance of the SETAR forecasting methods for an empirical example. In section 4.2 we analyse a later vintage of data for the period 1960.1-1994.4, and analyse forecasts of the period 1991.1-1994.4.

Only the MC, BS and SK methods of forecasting the SETAR models are considered. Based on the Monte Carlo results in section 3, the DE method has little to commend it, and whilst the NFEM generally performed well, the MC method is superior and simpler to implement (although computationally more intensive).

4.1. An empirical forecast comparison exercise for the period 1973:1-1990:4

The Tiao and Tsay (1994) study compared forecasts from a SETAR(2;2,2) to those from a linear AR model for US GNP. Following Tiao and Tsay, forecasts for the SETAR model are

calculated by estimating the model up to 1972.4, generating multi-step forecasts up to 12 steps ahead (by the MC method), and then re-estimating the model on an extended sample to 1973.1, etc., up to a final estimation period ending in 1987.4. This gives a sample of 60 sequences of 1 to 12 step ahead forecasts. However, certain parameters, namely $\{d, r, p_1, p_2\}$, were selected and held fixed by Tiao and Tsay at their full-sample values, rather than being those which would have been chosen based only on information available at the time the forecasts were made. Hence their study could be misleading as a guide to the genuine ex-ante performance of the SETAR model. Table 5 details the values of $\{d, r, p_1, p_2\}$ that would have been chosen on the basis of minimizing AIC for selected sub-samples.

Table 5

Sub-sample estimates of SETAR model for US GNP: 1947 to specified end date

End date	r	d	p ₁	α _{1,0}	α _{1,1}	α _{1,2}	p ₂	α _{2,0}	α _{2,1}
73:4	-0.008	3	1	1.049	0.329	-	1	0.369	0.443
74:4	-0.008	3	1	0.931	0.4150	-	1	0.316	0.460
75:4	-0.008	3	1	0.986	0.364	-	1	0.256	0.498
76:4	-0.008	3	1	0.986	0.364	-	1	0.269	0.492
77:4	-0.008	3	1	0.986	0.364	-	1	0.284	0.485
78:4	-0.087	3	1	1.078	0.195	-	1	0.315	0.494
79:4	-0.008	3	1	0.970	0.298	-	1	0.322	0.468
80:4	-0.008	2	2	-0.456	0.410	-0.918	1	0.659	0.328
81:4	-0.008	2	2	-0.533	0.415	-0.960	1	0.660	0.319
82:4	-0.008	2	2	-0.497	0.399	-0.850	1	0.611	0.353
83:4	-0.008	2	2	-0.479	0.396	-0.852	1	0.627	0.356
84:4	-0.058	2	2	-0.500	0.398	-0.868	1	0.614	0.363
85:4	-0.008	2	2	-0.479	0.396	-0.852	1	0.625	0.358
86:4	-0.008	2	2	-0.452	0.392	-0.838	1	0.621	0.351
87:4	-0.058	2	2	-0.469	0.394	-0.852	1	0.622	0.353

Figure 1 plots the MSFEs for multi-step ahead forecasts (1-12 steps) for the MC method relative to the AR(2), BS and SK methods, when we follow Tiao and Tsay and use a SETAR(2;2,2) model with d and r fixed at their full-sample values. We find gains to using the SETAR model (MC method) at short horizons. At 4-steps ahead the gain over the AR model is maximised at nearly 10%. At longer horizons there are no clear differences between the forecast methods for the SETAR, or between the SETAR and AR models.

Figure 1. Relative MSFEs for US GNP 1973.1-1990.4 for a SETAR(2;2,2), $r=0$, $d=2$
(The MC method MSFE is expressed relative to that for AR(2), BS and SK.)

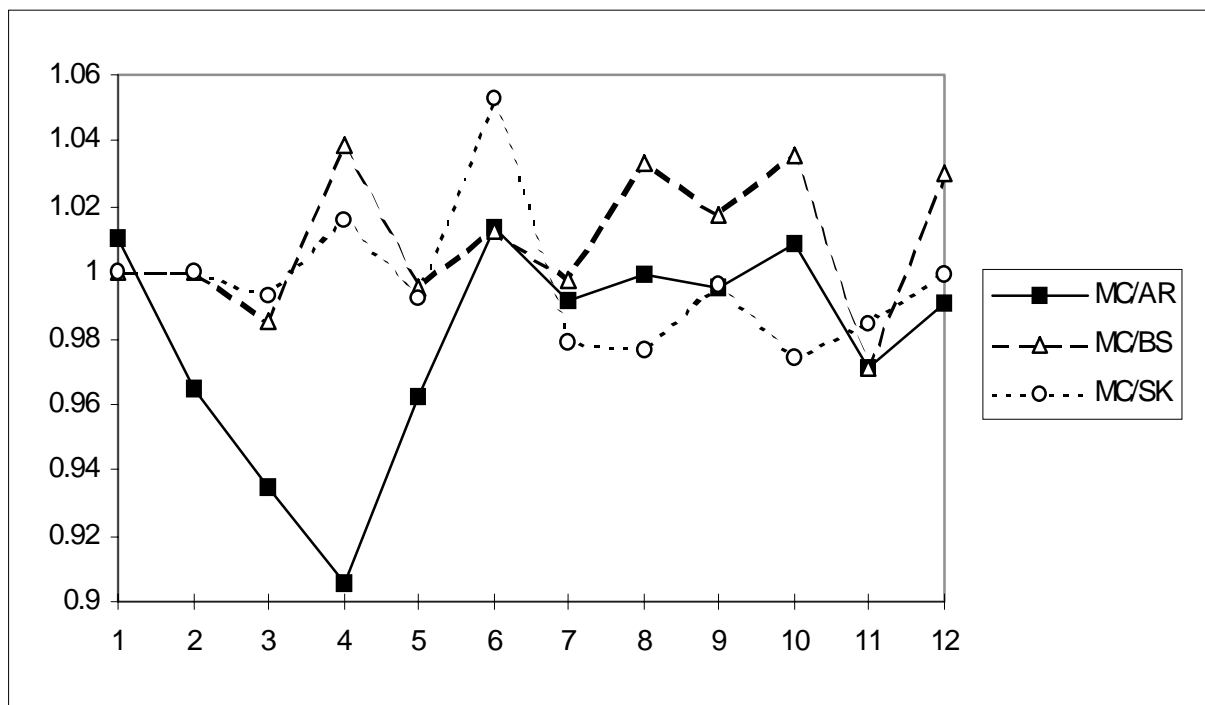
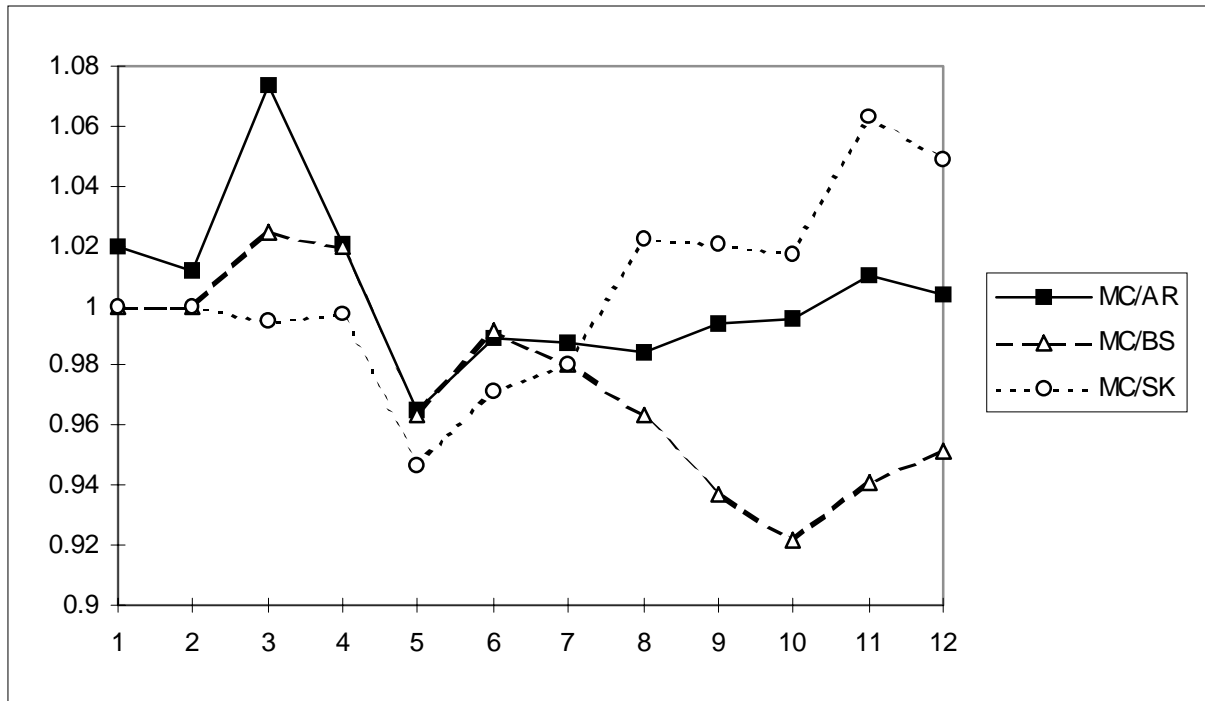


Figure 2 presents the results of a genuine ex ante forecast comparison with r , d , p_1 and p_2 all chosen to minimise AIC on each sub-sample (and correspond to the values reported in Table 5). We now find the three reported methods of forecasting the SETAR model yield inferior forecasts to the AR(2) model for horizons up to 3-steps ahead forecast. At 3-steps ahead, for

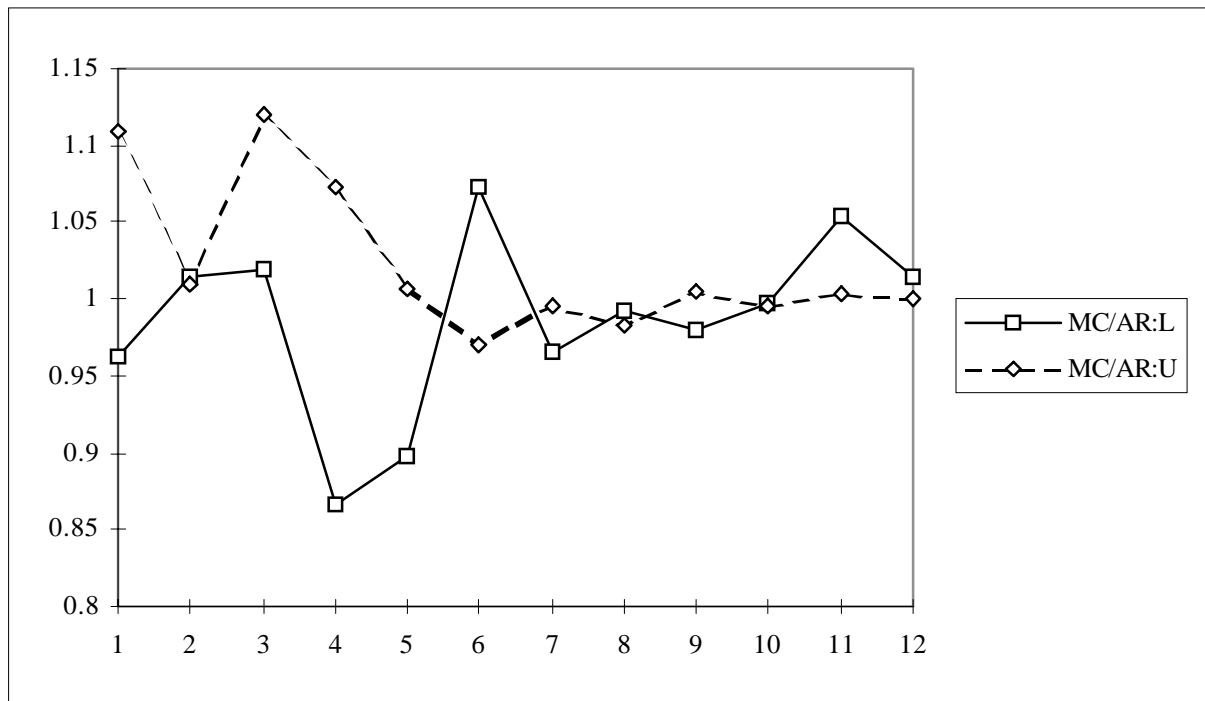
example, the MC method is approximately 7% worse than the AR(2). The best method at longer horizons appears to be the skeleton.

Figure 2. Relative MSFEs for US GNP 1973.1-1990.4 for a SETAR(2; p_1, p_2), r, d estimated
(The MC method MSFE is expressed relative to that for AR(2), BS and SK.)



However, a different story emerges if we condition on the regime. Figure 3 plots the relative MSFEs of the MC method to the AR(2) model conditional on the regime at the time the forecast was made. There are gains of 10% or more for the SETAR model at 4- and 5-steps ahead in the lower regime (MC/AR:L) with nothing thereafter, compared with substantial losses at all horizons less than 5-steps ahead in the upper regime (MC/AR:U).

Figure 3. Relative MSFEs for US GNP 1973.1-1990.4 for a SETAR(2; p_1, p_2), r, d estimated
 (The MC MSFE relative to AR(2) for the lower regime, L, and the upper regime, U)



These results are broadly in line with those of Clements and Smith (1996). They take the estimated SETAR(2;5,5) model of Potter (1995) and assume that this is the data generating process. In a Monte Carlo study they find that when the SETAR model is assumed known, then, using the MC method, there are gains of up to 16% at 1-step ahead and 21% at 2-steps ahead against an AR(2) model. However, when the AR coefficients and the lag orders are unknown (but r and d are known) the potential gains fall to only 1% at 1-step ahead and 5% at 2-steps ahead, when we do not condition upon being in a particular regime. These findings are also consistent with those of Hansen (1996). Hansen is unable to reject the linear model in favour of the SETAR for the period 1947-1990, suggesting that any gains to allowing for SETAR non-linearity are unlikely to be large.

4.2. An empirical forecast comparison exercise for the period 1991:1-1994:4

The second forecasting exercise is based on the more recent vintage of US GNP data at 1992 prices for the period 1960:1 to 1994:4. We calculate 12 sequences of 1- to 5-step ahead forecasts for the period 1991:1 to 1994:4, using a SETAR(2; p_1, p_2). The first forecast origin is 1990:4, and the last is 1993:3. On each estimation sub-sample $\{d, r, p_1, p_2\}$ are selected to minimize AIC in each regime (but not allowing holes in the lag distributions). A SETAR(2;1,1) is selected for each sub-sample, with $r \cong 0.2$ and $d=2$. Thus, omitting the late 1940s and 1950s from the estimation sample alters the specification of the SETAR model. This is not a data vintage effect – estimating the model over a similar sample period on the Potter dataset gave similar results, pointing to non-constancy of the SETAR model.

Nevertheless, our primary interest is in comparing the SETAR and AR models, and the alternative methods of generating multi-step forecasts for the former. The MSFEs from using the BS and SK methods for the SETAR, and the AR(2) and AR(3) models, all expressed relative to the MC method, are reported in Table 6. This table shows that the AR(2) model is 10% worse than the SETAR MC method at 1-step, similar at 2-steps, and nearly 20% better at 3-steps. The SK method is generally the best over the horizons considered.

The breakdown of the MSFEs into bias and variance components indicates that SETAR MC forecasts are more biased than those of the AR(2) at 1- through to 4-steps ahead, but have a smaller variance at 1-step ahead.

Table 6
Out-of-sample forecasts of US GNP 1991:1-1994:4

MSFEs relative to the MC method

Steps	AR(2)	AR(3)	BS	SK
1	1.10	1.26	1.00	1.00
2	1.01	1.19	0.91	0.92
3	0.82	0.87	0.93	0.89
4	1.08	1.05	1.58	0.86
5	1.03	1.11	1.98	0.92

Forecast biases

Steps	AR(2)	AR(3)	MC	BS	SK
1	0.04	0.07	0.07	0.07	0.07
2	-0.02	0.03	0.04	0.04	0.04
3	-0.03	0.03	0.06	-0.05	0.08
4	-0.07	-0.03	0.07	-0.10	0.08
5	-0.13	-0.10	-0.12	-0.23	0.02

Forecast error variances

Steps	AR(2)	AR(3)	MC	BS	SK
1	0.17	0.18	0.16	0.16	0.16
2	0.11	0.13	0.10	0.10	0.10
3	0.11	0.13	0.11	0.10	0.10
4	0.10	0.11	0.13	0.11	0.11
5	0.07	0.07	0.07	0.07	0.07

5. Conclusions

We have compared a number of methods that have been proposed in the literature for obtaining h -step ahead minimum mean square error forecasts for self-exciting threshold autoregressive (SETAR) models. The NFEM, BS and MC methods all perform reasonably well for a variety of different SETAR(2;1,1) models. The MC method is generally at least as good as that of the other methods we consider, and is simple to implement. The drawback of this method relative to an approximate analytical method such as NFEM is of course that it is relatively computer intensive. However, the computational cost of investigating the properties of this method by simulation (as in this paper) were not prohibitive given the current-generation fast computers.

When the SETAR model disturbances come from a highly asymmetric distribution, a Bootstrap method is to be preferred to the MC method. But even then, the costs to using the MC method in terms of forecast accuracy are relatively small, while the NFEM, based on the assumption of Gaussian disturbances, is less robust at long horizons.

The comparison of the accuracy of the forecasts from the SETAR to that of an AR model highlights those instances when a linear model provides a good approximation to the non-linear model, at least for forecasting. The sign and magnitude of the intercepts in the SETAR model were found to have a large effect on comparisons of the SETAR against the AR model forecasts.

In the empirical study we present multi-step ahead forecasts for US GNP from a SETAR and a linear AR model for two historical epochs. The SETAR forecasts are calculated using a number of the methods we study in the Monte Carlo. Our investigation casts some doubt on the robustness of SETAR models of US GNP over the two sample periods, and on the size of

the forecast accuracy gains that the SETAR models can deliver. We are able to replicate the finding in Tiao and Tsay (1994) that the SETAR model produces smaller MSFEs than an AR(2) model over the period 1973.1-1990.4, when certain parameters are set at their full-sample values. These gains disappear if the SETAR model is specified and estimated solely on the information available at the time the forecasts are made (as would be the case in a genuine forecasting exercise), but can be re-instated if forecasts are evaluated conditional upon the process being in a particular regime.

Finally, when analysing forecasts of the period 1991.1-1994.4 no clear picture emerges. There are unconditional (on the regime) gains to a SETAR model compared to an AR model at 1-step ahead, but not at 2-steps, and at 3-steps the ranking reverses.

References

- Al-Qassam, M. S. and J. A. Lane, 1989, Forecasting exponential autoregressive models of order 1, *Journal of Time Series Analysis*, 10, 95-113.
- Andrews, D. W. K., 1993, Tests for parameter instability and structural change with unknown change point, *Econometrica*, 61, 821-856.
- Clements, M. P. and D. F. Hendry, 1996, Multi-step estimation for forecasting, *Oxford Bulletin of Economics and Statistics*, 58, 657-684.
- Clements, M. P. and J. Smith, 1996, A Monte Carlo study of the forecasting performance of empirical SETAR models, Warwick Economic Research Papers No.464, Department of Economics, University of Warwick.
- De Gooijer, J. G. and K. Kumar, 1992, Some recent developments in non-linear time series modelling, testing and forecasting, *International Journal of Forecasting*, 8, 135-156.
- De Gooijer, J. G. and P. De Bruin, 1997, On SETAR forecasting, *Statistics and Probability Letters*, forthcoming.
- Diebold, F. X. and J. A. Nason, 1990, Nonparametric exchange rate prediction, *Journal of International Economics*, 28, 315-332.
- Granger, C. W. J. and T. Teräsvirta, 1993, *Modelling Nonlinear Economic Relationships*, Oxford University Press, Oxford.
- Hansen, B. E., 1996, Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica*, 64, 413-430.
- Lai, T. L. and G. Zhu, 1991, Adaptive prediction in non-linear autoregressive models and control systems, *Statistica Sinica*, 1, 309-334.

Potter, S., 1995, A nonlinear approach to U.S. GNP, *Journal of Applied Econometrics*, 10, 109-125.

Tiao, G. C. and R. S. Tsay, 1994, Some advances in non-linear and adaptive modelling in time-series, *Journal of Forecasting*, 13, 109-131.

Tong, H., 1978, On a threshold model, in C. H. Chen, ed, *Pattern Recognition and Signal Processing* (Sijhoff and Noordoff, Amsterdam),100-141.

Tong, H., 1995, *Non-Linear Time Series. A Dynamical System Approach*, Clarendon Press, Oxford.

Tong, H. and K. S. Lim, 1980, Threshold autoregression, limit cycles and cyclical data, *Journal of the Royal Statistical Society*, B 42, 245-292.