Journal of Forecasting
J. Forecast. **24**, 255–267 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/for.941

Detection of Regime Switches between Stationary and Nonstationary Processes and Economic Forecasting

KOSEI FUKUDA*
Nihon University, Japan

ABSTRACT
It often occurs that no model may be exactly right, and that different portions of the data may favour different models. The purpose of this paper is to propose a new procedure for the detection of regime switches between stationary and nonstationary processes in economic time series and to show its usefulness in economic forecasting. In the proposed procedure, time series observations are divided into several segments, and a stationary or nonstationary autoregressive model is fitted to each segment. The goodness of fit of the global model composed of these local models is evaluated using the corresponding information criterion, and the division which minimizes the information criterion defines the best model. Simulation and forecasting results show the efficacy and limitations of the proposed procedure. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS BIC; model selection; regime switch; structural change; unit root

INTRODUCTION

It is well known that economic systems exhibit occasional jumps from one regime to another. A nation's economy periodically switches from expansion to contraction and back again, and the dynamics differ between these two regimes. This feature is known as the asymmetry of business cycles, as advocated by Nefti (1984). A financial market periodically switches from a low-volatility regime to a high-volatility regime and back again. For example, this feature of volatility was modelled as a pure Markov-switching variance model by Turner *et al*. (1989). Thus the Markov switching model has been widely applied in time series econometrics literature, particularly since the seminal work of Hamilton (1989).

These regime switches may occur between stationary and nonstationary processes as well. For example, Ang and Bekaert (2002) pointed out that the US Federal Reserve tends to move short-term interest rates in a very persistent fashion during low-inflation periods. However, during high-inflation times, interest rate changes made by the US Federal Reserve become less persistent and have higher variance. This time series property makes it difficult to detect a unit root using the conventional hypothesis testing procedures. Nelson *et al*. (2001) investigated the poor performance of

* Correspondence to: Kosei Fukuda, College of Economics, Nihon University, 1-3-2 Misakicho, Chiyoda-ku, Tokyo 101-8360, Japan. E-mail: fukuda@eco.nihon-u.ac.jp

the Dickey–Fuller (1979, henceforth DF) type tests when the data undergo Markov regime switching.

In practice, there are many types of processes that have near unit roots, which are very difficult to distinguish from perfect unit root processes, given a finite sample. To examine these processes, two alternative approaches have been proposed. First, McCulloch and Tsay (1994) treated stationarity and nonstationarity as two competing models and allowed each observation to switch from one model to the other, with the transition being governed by a Markov process. For the monthly series of the industrial production index for the USA from 1947:1 to 1992:1, they concluded that there might be a structural change around 1971:12 and that the series switched from a nonstationary process to a stationary process. Second, Granger and Swanson (1997) introduced a class of stochastic unit root processes which have a root that is not constant, but is stochastic, and varies around unity. They compared forecasting performances of random walk with drift models, autoregressive models, time-varying parameter models and stochastic unit root models. Using monthly US macroeconomic time series, they concluded that stochastic unit root models are potentially useful in multi-step-ahead forecasting.

The purpose of this paper is to present a new procedure for the detection of regime switches between stationary and nonstationary processes and to show the usefulness of this procedure in forecasting economic time series. My procedure is an extended version of Kitagawa and Akaike (1978, hereafter referred to as KA). KA have proposed a procedure for fitting a locally stationary autoregressive model to a nonstationary time series using the Akaike information criterion (Akaike, 1974), AIC. In their procedure, time series observations are divided into several segments, and a stationary autoregressive model is fitted to each segment. The goodness of fit of the global model composed of these local stationary models is evaluated using the corresponding AIC, and the division which minimizes the AIC defines the best model. In the proposed procedure, however, not only a stationary model but also a nonstationary model for each segment is considered. For example, therefore, I examine the process which is derived by a stationary model in the first half and by a nonstationary model in the second half. In addition, not only the AIC but also the BIC, the Bayesian information criterion (Schwarz 1978), is introduced and compared.

Some comments on my definitions of a structural change and a regime switch used in this paper are appropriate here. It is natural to consider that the concept of structural change includes that of the regime switch. I therefore define a regime switch as only the case where a structural change occurs and the time series process changes from a stationary process to a nonstationary process, or *vice versa*. In addition, I use the terms 'structural break' and 'structural change' identically as the case may be.

The rest of the paper is organized as follows. In the next section, my procedure for the detection of regime switches between stationary and nonstationary processes is presented. In the third section, I examine the efficacy of my procedure using Monte Carlo simulations. In the fourth section, I apply the proposed procedure to 10 quarterly time series of industrial production indices of developed countries. In the fifth section, forecasting performances are compared using random walk with drift models and autoregressive models with unit root pretests allowing structural changes. Conclusions are presented in the last section.

MODEL SELECTION-BASED REGIME SWITCH DETECTION

In this section I describe a new procedure for the detection of regime switches between stationary and nonstationary processes, which is an extended version of KA. First, the KA procedure is reviewed briefly, and then how an extension is made to model such regime switches is discussed.

The KA procedure

KA considered the following autoregressive model with m structural changes (m + 1 regimes) for the observed time series y_t ($t = 1, \ldots, T$):

$$y_t = \phi_{i0} + \phi_{i1} y_{t-1} + \ldots + \phi_{ip_i} y_{t-p_i} + e_{it}, \quad t = T_{i-1} + 1, \ldots, T_i, \ i = 1, \ldots, m+1 \tag{1}$$

where e_{it} is the disturbance term generated from i.i.d. $N(0, \sigma_i^2)$, $T_0 = 0$ and $T_{m+1} = T$. All roots of $1 - \phi_{i1}L - \ldots - \phi_{iP_i}L^{P_i} = 0$ ($i = 1, \ldots, m + 1$) lie outside the unit circle. The indices ($T_1, \ldots, T_m$), or the breakpoints, are explicitly treated as unknown. The purpose is to estimate the unknown autoregressive coefficients together with the breakpoints when T observations on y_t are available. The goodness of fit of the global model composed of these local stationary models is evaluated using the corresponding AIC, and the division which minimizes the AIC defines the best model. KA's AIC is defined as:

$$\text{AIC} = \sum_{i=1}^{m+1} (T_i - T_{i-1}) \ln \hat{\sigma}_i^2 + 2 \sum_{i=1}^{m+1} (p_i + 2) \tag{2}$$

where $\hat{\sigma}_i^2$ is the maximum likelihood estimate for σ_i^2, and these residual variances are counted as parameters. In their minimum AIC procedure for the fitting of a locally stationary autoregressive model, the length of the basic span of data, K, is introduced. Therefore, the following conditions are imposed:

$$T_i - T_{i-1} = nK, \quad i = 1, \ldots, m+1 \tag{3}$$

where n is a positive integer. In addition, the maximum order of each autoregressive model, $p_{\max}$, must be set. As a numerical example, KA set $K = 100$ and $p_{\max} = 5$ for locally stationary processes with three regimes of $T = 900$.

The proposed procedure

I consider an extended version of the KA procedure with respect to the following three points. First, the regime switches between stationary and nonstationary processes are introduced in place of (1). Stationary and nonstationary processes are respectively described as

$$\begin{aligned} y_t &= \alpha_i + \beta_i t + z_t, \quad z_t = \phi_{i1} z_{t-1} + \ldots + \phi_{ip_i} z_{t-p_i} + e_{it}, \quad t = T_{i-1} + 1, \ldots, T_i \\ y_t &= \gamma_j + y_{t-1} + \phi_{j1} \Delta y_{t-1} + \ldots + \phi_{jp_j} \Delta y_{t-p_j} + e_{jt}, \quad t = T_{j-1} + 1, \ldots, T_j \end{aligned} \tag{4}$$

In the KA procedure, the stationary process around a time trend is not considered. In economic time series, however, this process seems to be more natural than a stationary process without a trend. The two-step procedure for a stationary process such as in (4) is advocated by Bhargava (1986), Perron (1989) and Schmidt and Phillips (1992).

Second, in place of a basic span, the minimum length of each segment, L, is introduced. Since the sample size of economic time series is generally limited, the notion of a basic span is not applicable. In place of KA's conditions (3), therefore, the following conditions are imposed:

$$T_i - T_{i-1} \geq L, \quad i = 1, \ldots, m+1$$

Third, breakpoints $T_1, \ldots, T_m$ should be treated as parameters to be estimated in the same way as in Yao (1988). In place of KA's definition (2), therefore, the following modified AIC (MAIC) is considered here:

$$\text{MAIC} = \sum_{i=1}^{m+1} (T_i - T_{i-1}) \ln \hat{\sigma}_i^2 + 2\left(m + \sum_{i=1}^{m+1} \theta_i\right)$$

where $\theta_i = 3 + p_i$ if stationary and $\theta_i = 2 + p_i$ if nonstationary.

In addition, the corresponding BIC is defined similarly as follows:

$$\text{MBIC} = \sum_{i=1}^{m+1} (T_i - T_{i-1}) \ln \hat{\sigma}_i^2 + \ln(T)\left(m + \sum_{i=1}^{m+1} \theta_i\right)$$

In the proposed procedure, all possible models are considered, and the best MAIC or MBIC model is selected among them. For example, in the case of $T = 100$, $m = 2$ and $L = 30$, 528 ($= 2 \times 2 \times 2 \times 66$) models are examined. The proposed procedure, which I call the information criterion-based model selection (ICBMS) procedure, is very computer-intensive. The development of a more efficient solution is left for future research.

SIMULATIONS

I examine the efficacy of my procedure via Monte Carlo simulations, and compare the results with those of the DF-type procedure. This section proceeds as follows. First, the data-generating processes (DGPs) considered in simulations are specified. Beginning with simple stationary and nonstationary processes, I consider several types of regime-switching processes. Second, the DF-type procedures are reviewed briefly, since it is reasonable to examine how this procedure performs in the case of regime switches, similarly to Nelson *et al*. (2001). Finally, I compare simulation results using my procedure with those using the DF-type procedure.

Data-generating processes

The following 13 DGPs are considered:

DGP 1 (S): $y_t = 0.4t + 1.4y_{t-1} - 0.6y_{t-2} + e_t, t = 1, \ldots, T$

DGP 2 (N): $y_t = 1.0 + 1.6y_{t-1} - 0.6y_{t-2} + e_t$

DGP 3 (SS): $y_t = 0.4t + 1.4y_{t-1} - 0.6y_{t-2} + e_t$ if $t \leq T_B$

$y_t = 0.4t - 0.2(t - T_B) + 1.4y_{t-1} - 0.6y_{t-2} + e_t$ if $t > T_B$

DGP 4 (SS): $y_t = y_{1t}$ if $t \leq T_B$, $y_t = y_{2t} + y_{1T_B} - y_{2T_B}$ if $t > T_B$

where $y_{1t} = 0.4t + 1.4y_{1t-1} - 0.6y_{1t-2} + e_t$

$y_{2t} = 0.8t + 0.6y_{2t-1} - 0.6y_{2t-2} + e_t$

DGP 5 (SN): $y_t = y_{1t}$ if $t \le T_B$, $y_t = y_{2t} + y_{1T_B} - y_{2T_B}$ if $t > T_B$
where $y_{1t} = 0.4t + 1.4y_{1t-1} - 0.6y_{1t-2} + e_t$
$y_{2t} = 0.5 + 1.6y_{2t-1} - 0.6y_{2t-2} + e_t$

DGP 6 (NN): $y_t = 1.0 + 1.6y_{t-1} - 0.6y_{t-2} + e_t$ if $t \le T_B$
$y_t = 0.5 + 1.6y_{t-1} - 0.6y_{t-2} + e_t$ if $t > T_B$

DGP 7 (NN): $y_t = y_{1t}$ if $t \le T_B$, $y_t = y_{2t} + y_{1T_B} - y_{2T_B}$ if $t > T_B$
where $y_{1t} = 1.0 + 1.6y_{1t-1} - 0.6y_{1t-2} + e_t$
$y_{2t} = 0.5 + 1.0y_{2t-1} + e_t$

DGP 8 (SSS): $y_t = 0.4t + 1.4y_{t-1} - 0.6y_{t-2} + e_t$ if $t \le T_{B1}$
$y_t = 0.4t - 0.2(t - T_{B1}) + 1.4y_{t-1} - 0.6y_{t-2} + e_t$ if $T_{B1} < t \le T_{B2}$
$y_t = 0.4t - 0.2(t - T_{B1}) - 0.2(t - T_{B2}) + 1.4y_{t-1} - 0.6y_{t-2} + e_t$ if $t > T_{B2}$

DGP 9 (SSS): $y_t = y_{1t}$ if $t \le T_{B1}$, $y_t = y_{2t} + y_{1T_{B1}} - y_{2T_{B1}}$ if $T_{B1} < t \le T_{B2}$
$y_t = y_{3t} + y_{2T_{B2}} - y_{3T_{B2}}$ if $t > T_{B2}$
where $y_{1t} = 0.4t + 1.4y_{1t-1} - 0.6y_{1t-2} + e_t$
$y_{2t} = 0.8t + 0.6y_{2t-1} - 0.6y_{2t-2} + e_t$
$y_{3t} = 1.2t - 0.6y_{3t-2} + e_t$

DGP 10 (SNS): $y_t = y_{1t}$ if $t \le T_{B1}$, $y_t = y_{2t} + y_{1T_{B1}} - y_{2T_{B1}}$ if $T_{B1} < t \le T_{B2}$
$y_t = y_{3t} + y_{2T_{B2}} - y_{3T_{B2}}$ if $t > T_{B2}$
where $y_{1t} = 0.4t + 1.4y_{1t-1} - 0.6y_{1t-2} + e_t$
$y_{2t} = 0.5 + 1.0y_{2t-1} + e_t$
$y_{3t} = 0.4t + 0.6y_{3t-1} - 0.6y_{3t-2} + e_t$

DGP 11 (NSN): $y_t = y_{1t}$ if $t \le T_{B1}$, $y_t = y_{2t} + y_{1T_{B1}} - y_{2T_{B1}}$ if $T_{B1} < t \le T_{B2}$
$y_t = y_{3t} + y_{2T_{B2}} - y_{3T_{B2}}$ if $t > T_{B2}$
where $y_{1t} = 0.5 + 1.6y_{1t-1} - 0.6y_{1t-2} + e_t$
$y_{2t} = 0.4t + 1.4y_{2t-1} - 0.6y_{2t-2} + e_t$
$y_{3t} = 1.0 + 1.0y_{3t-1} + e_t$

DGP 12 (NNN): $y_t = 1.0 + 1.6y_{t-1} - 0.6y_{t-2} + e_t$ if $t \le T_{B1}$
$y_t = 0.5 + 1.6y_{t-1} - 0.6y_{t-2} + e_t$ if $T_{B1} < t \le T_{B2}$
$y_t = 1.6y_{t-1} - 0.6y_{t-2} + e_t$ if $t > T_{B2}$

DGP 13 (NNN): $y_t = y_{1t}$ if $t \le T_{B1}$, $y_t = y_{2t} + y_{1T_{B1}} - y_{2T_{B1}}$ if $T_{B1} < t \le T_{B2}$
$y_t = y_{3t} + y_{2T_{B2}} - y_{3T_{B2}}$ if $t > T_{B2}$
where $y_{1t} = 0.5 + 1.8y_{1t-1} - 0.8y_{1t-2} + e_t$
$y_{2t} = 1.0 + 1.5y_{2t-1} - 0.5y_{2t-2} + e_t$
$y_{3t} = 1.5 + 1.2y_{3t-1} - 0.2y_{3t-2} + e_t$

In each DGP above, the letter within the parentheses denotes 'stationary' or 'nonstationary'. For example, 'NSN' within the parentheses of DGP 11 indicates that this series has three regimes, or two breaks, with nonstationary processes for the first and third regimes and a stationary process for the second regime. The conventional DGPs when considering the unit root test allowing for structural change are 1, 2, 3 and 8. The DGPs 4, 6, 7, 9, 12 and 13 are beyond the scope of conventional unit root and structural change literature. The DGPs 5, 10 and 11 are specified in order to examine the properties of regime switches. In the above DGPs, e_t is generated from n.i.i.d.(0, 1), $T_B = T/2$, $T_{B1} = T/3$, $T_{B2} = 2T/3$. In each simulation, I set $T = 120$ and $p_{\max} = 4$. The maximum number of regime switches or structural changes, $m_{\max}$, is fixed as $m_{\max} = 2$. The number of replications is 300.

The Dickey–Fuller-type tests

The DF-type test employed here is the same as that proposed by Ohara (1999). Ohara developed an extended version of the Zivot and Andrews (1992) (henceforth ZA) test. He considered the following DF-type test regression allowing q trend breaks:

$$y_t = a + b_0 t + b_1 DT_t^*(\lambda_1) + b_2 DT_t^*(\lambda_2) + \ldots + b_q DT_t^*(\lambda_q) + c y_{t-1} + \sum_{j=1}^{k} d_j \Delta y_{t-j} + e_t$$

where $0 < \lambda_1 < \ldots < \lambda_q < 1$, $DT_t^*(\lambda_i) = t - T_{Bi}$ if $t > T_{Bi}$ and $DT_t^*(\lambda_i) = 0$ otherwise. T_{Bi} denotes the ith breakpoint, and λ_i denotes the break fraction such as $\lambda_i = T_{Bi}/T$. The null hypothesis is

$$c = 1, \quad (b_0, b_1, b_2, \ldots, b_q) = (0, 0, 0, \ldots, 0)$$

The alternative hypothesis is

$$c < 1, \quad (b_0, b_1, b_2, \ldots, b_q) \neq (0, 0, 0, \ldots, 0)$$

Ohara's technique is the same as that of ZA, and therefore he obtains the minimum of the sequence of unit root test statistics, by sequentially incrementing the break fractions. He derived the asymptotic distributions for this minimum under the preceding null hypothesis and tabulated the critical values. Ohara's test includes DF and ZA tests as special cases of $q = 0$ and $q = 1$, respectively.

Some comments on lag length selection and on trimming are appropriate here. The number of extra regressors, k, is determined using the same selection procedure as that used by Perron (1989); that is, working backward from $k = k_{\max}$, I chose the first value of k such that the t-statistic on $\hat{d}_k$ was greater than 1.6 in absolute value and the t-statistic on $\hat{d}_l$ for $l > k$ was less than 1.6 in absolute value. I set $k_{\max} = 4$.

In monitoring multiple structural changes, I cannot consider change points too close to the beginning or end of the sample or to other change points. Let δ denote the trimming value. The location of structural changes is restricted as

$$\delta < T_{B1} < T_{B1} + \delta < T_{B2} < T - \delta$$

where I set $\delta = 5$ as in Ohara.

Comparisons between the DFZAO and ICBMS procedures

In the DF-type procedure, the following steps are taken. First, the DF test is carried out. If the null hypothesis can be rejected at the 5% level, then the stationary process with no break (hereafter referred to as SB0) is selected. If the null hypothesis cannot be rejected, the ZA test is next implemented. If the null hypothesis can be rejected, then the stationary process with one break (SB1) is selected. Otherwise, Ohara's test is finally employed. If the null hypothesis can be rejected, then the stationary process with two breaks (SB2) is selected. Otherwise, the nonstationary process with no break (NB0) is selected. Hereafter, I call this selection procedure the DFZAO procedure.

In the ICBMS procedure, on the other hand, all possible models are considered and estimated, and related IC values are stored. Next, the best IC model is selected among alternative models. It should be noted that the NB1 or NB2 process can be selected by the ICBMS procedure but not by the DFZAO procedure by definition.

In the ICBMS procedure, the minimum length of each segment is fixed as $L = 26$. There is no strictly economic or statistical reason for this setting. In the subsequent studies, I apply this procedure to quarterly economic time series. The average of the period of one business cycle in postwar USA is 22 quarters, and I believe that the minimum data length for estimating an autoregressive model needs more than the period of one business cycle. Finally, given the maximum number of structural changes, such as $m_{\max} = 2$, the maximum order of each autoregressive model, such as $p_{\max} = 4$, and the minimum data length, such as $L = 26$, the best AIC and BIC models are selected among all possible models.

Simulation results

Table I shows the frequency counts of selected processes. First, the results of the DFZAO procedure are examined. For DGPs 1, 3 and 8, expected results are obtained. For DGP 2, however, in the terminology of hypothesis testing, size distortions occur, as expected by Ohara. The true model (in this case, NB0) is selected only at 65%. Ohara reported, for example, that the 5% critical value obtained using the asymptotic distribution is −5.35, but that determined using the finite sample distribution ($T = 150$) is −5.89. For other DGPs, the results are more mixed, as expected on the basis of the condition that these DGPs are not considered in the DFZAO procedure. For DGP 5, for example, where the regime switch from a stationary process to a nonstationary process occurs, frequency counts are separated into three processes, except for SB0. In another case of DGP 7, SB1 is incorrectly selected at 73%. The AIC-based model selection (henceforth, AICBMS) procedure performs poorly. The results suggest a consistent bias towards the selection of stationary processes with more structural changes or regime switches.

On the other hand, the BIC-based model selection (henceforth, BICBMS) procedure performs better than the AICBMS procedure, but shows a consistent bias towards the selection of a nonstationary process with less structural changes or regime switches. As a typical case, no true model (SSS) is selected for DGP 8, whereas the true model is selected at 73% using the AICBMS procedure. This is because, in the true DGP, the parameters for three autoregressive processes are the same, whereas they are assumed to be different in the BICBMS procedure. Thus, more parsimonious models 'N' and 'NN' are selected at 32% and 42%, respectively. As discussed by Hansen (2001), the need for two structural breaks also reduces the distinction between the trend break and random walk models. Another reason why a nonstationary autoregressive model is preferred to a stationary model in BIC is that the former is more parsimonious than the latter. As a simple example, consider a persistent time series such as unemployment rate or interest rate. It often occurs that a random walk model shows as good a performance as a stationary AR(1) model. In this case, the former is more parsimonious, by one parameter, than the latter.

Table I. Frequency counts of selected processes

(A) Dickey–Fuller, Zivot–Andrews and Ohara using the 5% significance level

	SB0	SB1	SB2	NB0
DGP 1	0.97	0.01	0.01	0.00
DGP 2	0.08	0.15	0.12	0.65
DGP 3	0.00	0.99	0.01	0.00
DGP 4	0.00	1.00	0.00	0.00
DGP 5	0.03	0.40	0.34	0.24
DGP 6	0.01	0.27	0.15	0.57
DGP 7	0.01	0.73	0.17	0.09
DGP 8	0.00	0.01	0.94	0.05
DGP 9	0.17	0.83	0.00	0.00
DGP 10	0.12	0.72	0.11	0.05
DGP 11	0.00	0.19	0.53	0.29
DGP 12	0.01	0.17	0.20	0.63
DGP 13	0.17	0.26	0.18	0.40

(B) AIC- and BIC-based model selections

		S	N	SS	NS	SN	NN	SSS	NSS	SNS	NNS	SSN	NSN	SNN	NNN
DGP 1	AIC	0.19	0.00	0.30	0.02	0.02	0.00	0.35	0.02	0.04	0.00	0.04	0.01	0.00	0.00
	BIC	0.93	0.02	0.00	0.02	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DGP 2	AIC	0.02	0.00	0.13	0.02	0.03	0.01	0.24	0.11	0.08	0.05	0.15	0.07	0.06	0.03
	BIC	0.03	0.79	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
DGP 3	AIC	0.00	0.00	0.23	0.01	0.00	0.00	0.55	0.03	0.10	0.01	0.05	0.01	0.01	0.00
	BIC	0.00	0.39	0.18	0.10	0.11	0.15	0.00	0.00	0.02	0.01	0.00	0.00	0.01	0.02
DGP 4	AIC	0.00	0.00	0.34	0.01	0.00	0.00	0.49	0.04	0.10	0.01	0.01	0.00	0.00	0.00
	BIC	0.00	0.00	0.50	0.43	0.01	0.02	0.00	0.00	0.01	0.02	0.00	0.00	0.01	0.00
DGP 5	AIC	0.00	0.00	0.11	0.00	0.02	0.00	0.37	0.02	0.12	0.02	0.21	0.02	0.11	0.00
	BIC	0.00	0.51	0.01	0.01	0.21	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.04
DGP 6	AIC	0.00	0.00	0.06	0.02	0.01	0.02	0.24	0.12	0.11	0.04	0.15	0.07	0.09	0.06
	BIC	0.00	0.51	0.00	0.01	0.01	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06
DGP 7	AIC	0.00	0.00	0.09	0.03	0.01	0.01	0.23	0.14	0.13	0.08	0.13	0.07	0.04	0.03
	BIC	0.00	0.00	0.00	0.01	0.03	0.77	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.16
DGP 8	AIC	0.00	0.00	0.00	0.00	0.00	0.00	0.73	0.04	0.11	0.01	0.07	0.02	0.01	0.00
	BIC	0.00	0.32	0.00	0.04	0.07	0.42	0.00	0.00	0.00	0.03	0.01	0.02	0.02	0.07
DGP 9	AIC	0.00	0.00	0.00	0.00	0.00	0.00	0.89	0.06	0.03	0.00	0.00	0.01	0.00	0.00
	BIC	0.00	0.00	0.10	0.34	0.01	0.03	0.12	0.23	0.02	0.08	0.01	0.03	0.01	0.01
DGP 10	AIC	0.00	0.00	0.01	0.00	0.00	0.00	0.70	0.06	0.19	0.03	0.01	0.00	0.00	0.00
	BIC	0.00	0.00	0.03	0.09	0.01	0.08	0.00	0.01	0.16	0.49	0.00	0.01	0.04	0.08
DGP 11	AIC	0.00	0.00	0.03	0.01	0.01	0.00	0.39	0.21	0.06	0.01	0.16	0.10	0.02	0.01
	BIC	0.00	0.10	0.00	0.02	0.04	0.61	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.18
DGP 12	AIC	0.00	0.00	0.04	0.01	0.01	0.00	0.23	0.13	0.11	0.06	0.15	0.12	0.08	0.05
	BIC	0.00	0.27	0.00	0.00	0.00	0.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14
DGP 13	AIC	0.01	0.00	0.08	0.01	0.04	0.01	0.22	0.15	0.12	0.06	0.14	0.09	0.05	0.03
	BIC	0.01	0.25	0.00	0.01	0.01	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.10

Notes: The symbol 'S' indicates 'stationary', 'N' indicates 'nonstationary' and 'B' indicates 'break'. For example, SB0 indicates that this series is generated from a stationary process with no break. Similarly, NSN indicates that this series has three regimes, or two breaks, with nonstationary processes for the first and third regimes and a stationary process for the second regime. $T = 120$; 300 replications.

APPLICATIONS

In this section, I apply the DFZAO and ICBMS procedures to 10 quarterly time series of industrial production indices of developed countries: USA, Canada, Australia, Japan, Finland, France, Germany, Ireland, Netherlands, UK. The data set is obtained from International Financial Statistics. Each series is seasonally adjusted, and multiplied by 100 after log transformation for the period 1957 I–2002 III.

The following conditions are imposed: $k_{\max} = 12$, as suggested by ZA and Ohara; $m_{\max} = 2$ and $p_{\max} = 6$, on account of computational burden; and $L = 28$, as suggested in the preceding section.

Table II shows estimation results. The results of applying the DFZAO procedure suggest that five series (USA, Canada, France, Germany, Netherlands) are stationary with one break, one series (Japan) is stationary with two breaks, and four series (Australia, Finland, Ireland, UK) are nonstationary without a break. Using the AICBMS procedure, the three-regime model is selected for all series, and four series are stationary for all regimes, as expected on the basis of simulation results in the preceding section. In contrast, the BICBMS procedure has a tendency to select a nonstationary process. The three-regime model with three nonstationary processes (NNN) is selected for five series (USA, Germany, Ireland, Netherlands, UK). The two-regime model with two nonstationary processes (NN) is selected for three series (Canada, Finland, France). The crucial difference in estimation results between DFZAO and ICBMS procedures is on the breakpoints. However, I cannot conclude which result is correct, since the true DGP is unknown.

COMPARISONS OF FORECASTING PERFORMANCES

One clear way to judge the relevance of a model is to ask how well it performs compared to other models in out-of-sample forecasting using actual data. In this section, the forecasting performances of ICBMS procedures are compared to other forecasting procedures using the 10 quarterly time series analysed above. The models considered here are random walk with drift (henceforth RWD) models, and autoregressive models with unit root pretests allowing structural changes. Stock and Watson (1999) compared forecasting performances using 215 US monthly macroeconomic time series and 49 univariate forecasting methods, including autoregressions, exponential smoothing, artificial neural networks and smooth transition autoregressions, and concluded that the best overall performance of a single method is achieved using autoregression with unit root pretests.

In autoregressive models with unit root pretests allowing structural changes, the DFZAO procedure is implemented. As a result of model selection using the DFZAO procedure, one forecasting model is obtained for each series. For example, if a stationary model without a break is selected using the DF test, forecasting using this model is employed. If a stationary model with two breaks is not selected using Ohara's test, forecasting using a nonstationary model is implemented. The conditions imposed are the same as in the preceding section.

The reason for not considering nonlinear models is as follows. Engle (1994) considered a Markov switching model and found that it does not forecast exchange rates better than a random walk. Similar results were obtained in forecasting the US unemployment rates (Montgomery *et al*., 1998). In addition, the McCulloch and Tsay (1994) method is not applicable to forecasting, because their model allows each observation to switch from one model to the other. Their method is suitable not for forecasting but for outlier detection. Finally, stochastic unit root models developed by Granger and

Table II. Estimation results

	Dickey–Fuller, Zivot–Andrews and Ohara						AIC-based model selection							BIC-based model selection						
	$\hat{T}_{B1}$	$\hat{T}_{B2}$	k	$\hat{c}$	$\hat{t}_c$	$S(\hat{e})$	$\hat{T}_{B1}$	$\hat{T}_{B2}$	Regime	Lag	$S(\hat{e}_1)$	$S(\hat{e}_2)$	$S(\hat{e}_3)$	$\hat{T}_{B1}$	$\hat{T}_{B2}$	Regime	Lag	$S(\hat{e}_1)$	$S(\hat{e}_2)$	$S(\hat{e}_3)$
USA	66 I		10	−0.117	−4.72*	1.242	73 III	81 III	SSS	L622	1.638	1.410	0.719	73 IV	81 IV	NNN	L101	1.804	1.824	0.716
Canada	72 IV		10	−0.187	−5.22**	1.391	74 I	83 IV	SSS	L232	1.188	1.686	1.089	83 III		NN	L11	1.605	1.127	
Australia	73 II	77 IV	9	−0.318	−5.22	1.634	71 III	88 IV	SSS	L221	1.421	1.905	1.126	71 IV	88 III	SNN	L210	1.466	1.935	1.218
Japan	71 IV	91 I	11	−0.277	−6.55**	1.345	73 IV	92 III	SSS	L253	1.359	1.041	1.481	73 IV	92 III	NNS	L113	1.474	1.200	1.481
Finland	72 IV	95 III	9	−0.178	−4.30	2.260	70 III	90 IV	NSS	L564	2.315	1.951	1.759	70 III		NN	L40	2.340	2.282	
France	73 II		12	−0.186	−4.59*	1.398	66 II	74 II	SSN	L400	1.401	0.939	1.232	74 II		NN	L00	1.606	1.232	
Germany	72 I		11	−0.221	−4.77*	1.512	83 II	91 I	NSS	L463	1.915	0.909	1.052	66 III	82 IV	NNN	L010	2.066	1.572	1.391
Ireland	68 II	90 I	11	−0.298	−4.32	2.611	86 IV	94 II	SSN	L122	2.440	1.552	3.797	86 II	93 IV	NNN	L000	2.545	1.525	4.229
Netherlands	71 I		11	−0.142	−4.52*	1.725	74 III	95 III	NSS	L455	1.330	1.862	0.461	74 III	95 II	NNN	L010	1.373	2.007	0.764
UK	72 III	79 IV	8	−0.224	−4.19	1.645	73 IV	84 I	SSN	L422	1.568	1.618	0.834	73 IV	83 IV	NNN	L001	1.764	1.860	0.839

Notes: Regression: $\Delta y_t = \hat{a} + \hat{b}_0 t + \hat{b}_1 DT_t^*(\hat{\lambda}_1) + \hat{b}_2 DT_t^*(\hat{\lambda}_2) + \hat{c} y_{t-1} + \sum_{t=1}^{k} \hat{d}_i \Delta y_{t-i} + \hat{e}_t$. The symbols * and ** denote statistical significance at the 5% and 1% levels, respectively. The column 'Regime' shows the regime switching process. In this column, 'S' denotes 'stationary' and 'N' denotes 'nonstationary'. For example, SNS indicates that this series has three regimes, or two breaks, with stationary processes for the first and third regimes and a nonstationary process for the second regime. The column 'Lag' shows the lag selection process. For example, L40 indicates that this series has two regimes with four lags for the first regime and no lag for the second regime. T_{Bi} denotes the ith breakpoint.

Swanson (1997) do not perform as well as random walk with drift models in forecasting monthly US macroeconomic time series at one- and five-step-ahead.

All forecasts simulate real-time implementation, that is, they are fully recursive (e.g., Montgomery *et al*., 1998; Stock and Watson, 1999). For example, the case of a four-step-ahead forecast is considered as follows. First, the estimation is carried out for the period 1957 I–1976 IV and the forecast for 1977 IV is performed. Next, the estimation is carried out for the period 1957 I–1977 I and the forecast for 1978 I is performed. The same implementation is continued 100 times in total. The last estimation is therefore carried out for the period 1957 I–2001 III and the forecast for 2002 III is performed.

Table III shows comparisons of forecast performances at one- to four-step-ahead. The mean square forecast error (MSFE) and mean forecast error (MFE) are considered in order to evaluate forecast performances. Except for RWD models for Japan, MFEs are relatively small. Since in Japan growth rates after the oil price shock were less than half those before the shock, large negative bias (overprediction) occurs when using RWD models. Thus I focus my discussion on the MSFE for the other nine series shown in the table. First, the DFZAO procedure performs worse than other procedures. Second, the BICBMS procedure slightly outperforms other procedures. For six series (USA, Australia, Finland, Germany, Ireland, UK), forecasting performances of the BICBMS and RWD procedures are almost the same and better than those of the AICBMS procedure. The BICBMS procedure outperforms the RWD procedure for two series (France, Netherlands), and the converse result is obtained for one series (Canada).

CONCLUSIONS

In this paper I proposed a new procedure for the detection of regime switches between stationary and nonstationary processes in economic time series. In the proposed procedure, time series observations are divided into several segments, and a stationary or nonstationary autoregressive model is fitted to each segment. The goodness of fit of the global model composed of these local models is measured by the corresponding information criterion, and the division which minimizes the information criterion defines the best model. Simulation results show that the BIC-based model selection procedure performs better than the AIC-based procedure, but shows a consistent bias towards the selection of a nonstationary process with less structural changes or regime switches. Forecasting performances were compared between alternative models such as random walk with drift models and autoregressive models with unit root pretests allowing structural changes. The results showed that the BIC-based procedure slightly outperforms other procedures.

In practice, it often occurs that neither model may be exactly right, and that the unknown structure of the data may change over time. Thus different portions of the data may favour different models. Since a conventional unit root test employs only a global summary statistic, it is not applicable to this situation. The proposed method, on the other hand, can monitor the evolution of the model structure.

In examining a data set, a time series analyst usually asks which of several competing models best fits all of the data. I relax this requirement here for the problem of choosing between a stationary model and a nonstationary model. A great advantage of my procedure, compared to other procedures, is the consistent model evaluation. In the proposed procedure, for example, a partially explosive autoregressive process and/or moving average error term can easily be introduced, and the efficacy of these model changes can be consistently evaluated using the information criterion.

Table III. Comparisons of forecast performances

Country	Procedure	Mean square forecast error				Mean forecast error			
		One-step-ahead	Two-step-ahead	Three-step-ahead	Four-step-ahead	One-step-ahead	Two-step-ahead	Three-step-ahead	Four-step-ahead
USA	RWD	1.84	5.64	10.20	15.40	–0.17	–0.35	–0.55	–0.76
	DFZAO	1.73	7.04	15.20	25.74	–0.01	–0.03	–0.03	–0.03
	AICBMS	1.63	6.22	12.99	21.20	–0.24	–0.58	–0.96	–1.30
	BICBMS	1.49	4.97	9.85	16.17	–0.12	–0.28	–0.46	–0.65
Canada	RWD	3.61	10.02	18.74	29.23	–0.43	–0.84	–1.25	–1.64
	DFZAO	3.89	12.39	25.27	45.11	–0.02	–0.04	–0.03	0.00
	AICBMS	4.15	12.11	24.82	44.78	0.02	0.13	0.35	0.67
	BICBMS	3.52	10.82	22.41	39.17	–0.12	–0.23	–0.26	–0.27
Australia	RWD	2.84	6.10	9.51	12.77	–0.30	–0.61	–0.92	–1.24
	DFZAO	4.06	10.71	19.64	29.49	0.18	0.42	0.65	0.82
	AICBMS	3.33	7.76	12.72	17.69	0.01	0.08	0.15	0.20
	BICBMS	2.93	6.50	10.26	14.21	0.21	0.40	0.59	0.77
Japan	RWD	4.92	16.61	34.32	55.80	–1.41	–2.83	–4.21	–5.56
	DFZAO	2.66	9.09	21.16	37.29	0.20	0.50	0.83	1.18
	AICBMS	2.68	8.57	20.83	37.04	–0.18	–0.45	–0.78	–1.18
	BICBMS	2.66	7.99	17.51	28.05	–0.29	–0.75	–1.22	–1.72
Finland	RWD	4.18	10.39	16.76	24.73	–0.25	–0.51	–0.70	–0.88
	DFZAO	6.31	16.36	25.83	39.22	0.11	0.25	0.45	0.64
	AICBMS	4.73	11.72	19.18	28.47	0.15	0.33	0.54	0.73
	BICBMS	4.18	10.30	16.94	25.86	0.06	0.09	0.21	0.32
France	RWD	1.91	4.43	7.99	12.20	–0.53	–1.07	–1.60	–2.11
	DFZAO	2.67	6.16	11.14	16.85	–0.06	–0.11	–0.16	–0.17
	AICBMS	1.97	3.54	5.56	7.49	0.05	0.02	0.01	0.00
	BICBMS	1.73	3.68	6.36	9.27	–0.13	–0.29	–0.42	–0.53
Germany	RWD	2.97	6.75	11.65	17.76	–0.55	–1.09	–1.64	–2.17
	DFZAO	3.30	7.62	12.74	19.81	–0.15	–0.30	–0.45	–0.61
	AICBMS	3.40	7.57	12.72	19.95	–0.15	–0.30	–0.51	–0.74
	BICBMS	3.26	7.16	10.88	16.31	–0.06	–0.17	–0.28	–0.40
Ireland	RWD	8.22	17.96	26.07	34.62	0.26	0.64	0.98	1.23
	DFZAO	9.76	22.30	36.31	50.05	0.26	0.64	1.07	1.48
	AICBMS	9.61	21.08	30.71	41.43	0.21	0.46	0.65	0.85
	BICBMS	8.12	17.60	25.50	33.08	0.27	0.65	1.00	1.26
Netherlands	RWD	4.23	7.29	10.88	15.64	–0.65	–1.31	–1.97	–2.63
	DFZAO	4.52	7.86	11.97	16.45	0.29	0.53	0.75	1.03
	AICBMS	4.20	6.17	7.90	10.44	–0.16	–0.27	–0.43	–0.54
	BICBMS	4.06	6.21	7.95	11.33	0.06	–0.02	–0.08	–0.15
UK	RWD	1.97	5.37	9.37	13.62	–0.24	–0.51	–0.78	–1.03
	DFZAO	3.46	10.49	18.93	29.20	–0.45	–0.90	–1.30	–1.67
	AICBMS	2.21	6.14	10.58	15.37	–0.20	–0.46	–0.73	–0.99
	BICBMS	1.96	5.27	9.03	13.01	–0.12	–0.29	–0.45	–0.61

Notes: RWD denotes random walk with drift, DFZAO denotes Dickey–Fuller, Zivot–Andrews and Ohara, and AICBMS and BICBMS denote AIC- and BIC-based model selections, respectively.

REFERENCES

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **AC-19**: 716–723.

Ang A, Bekaert G. 2002. Regime switches in interest rates. *Journal of Business & Economic Statistics* **20**: 163–182.

Bhargava A. 1986. On the theory of testing for unit roots in observed time series. *Review of Economic Studies* **52**: 369–384.

Dickey DA, Fuller WA. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74**: 427–431.

Engle C. 1994. Can the Markov switching model forecast exchange rates? *Journal of International Economics* **36**: 151–165.

Granger CWJ, Swanson NR. 1997. An introduction to stochastic unit-root process. *Journal of Econometrics* **80**: 35–62.

Hamilton JD. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**: 357–384.

Hansen BE. 2001. The new econometrics of structural change: dating breaks in U.S. labor productivity. *Journal of Economic Perspective* **15**: 117–128.

Kitagawa G, Akaike H. 1978. A procedure for the modeling of non-stationary time series. *Annals of the Institute of Statistical Mathematics* **30B**: 351–363.

McCulloch RE, Tsay RS. 1994. Bayesian inference of trend- and difference-stationarity. *Econometric Theory* **15**: 523–539.

Montgomery AL, Zarnowitz V, Tsay RS, Tiao GC. 1998. Forecasting the U.S. unemployment rate. *Journal of the American Statistical Association* **93**: 478–493.

Nefti SN. 1984. Are economic time series asymmetric over business cycle? *Journal of Political Economy* **92**: 307–328.

Nelson CR, Piger J, Zivot E. 2001. Markov regime switching and unit-root tests. *Journal of Business & Economic Statistics* **19**: 404–415.

Ohara H. 1999. A unit root test with multiple trend breaks: a theory and an application to US and Japanese macroeconomic time series. *Japanese Economic Review* **50**: 266–290.

Perron P. 1989. The great crash, the oil-price shock, and the unit-root hypothesis. *Econometrica* **57**: 1361–1402.

Schmidt P, Phillips PCB. 1992. LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics* **54**: 257–287.

Schwarz G. 1978. Estimating the dimension of a model. *The Annals of Statistics* **6**: 461–464.

Stock JH, Watson MW. 1999. A comparison of linear and nonlinear univariate time series models for forecasting macroeconomic time series. In *Cointegration, Causality, and Forecasting*, Engle RF, White H (eds). Oxford University Press: Oxford.

Turner CM, Startz R, Nelson CR. 1989. A Markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics* **25**: 3–22.

Yao YC. 1988. Estimating the number of change-points via Schwarz' criterion. *Statistics and Probability Letters* **6**: 181–189.

Zivot E, Andrews DWK. 1992. Further evidence on the great crash, the oil price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics* **10**: 251–270.

Author's biography:
Kosei Fukuda is Associate Professor of Economics at Nihon University, Tokyo, Japan. He served as an economist in the Economic Planning Agency of the Japanese government (1986–2000). His research interests include time series econometrics and forecasting.

Author's address:
Kosei Fukuda, College of Economics, Nihon University, 1-3-2 Misakicho, Chiyoda-ku, Tokyo 101-8360, Japan.

