OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 66, 5 (2004) 0305-9049

Tests for Stationarity in Series with Endogenously Determined Structural Change*

DAVID I. HARVEY† and TERENCE C. MILLS‡

†*School of Economics, University of Nottingham, University Park, Nottingham, UK (e-mail: dave.harvey@nottingham.ac.uk)*
‡*Department of Economics, Loughborough University, Loughborough, Leicestershire, UK (e-mail: t.c.mills@lboro.ac.uk)*

Abstract

We consider tests of the null hypothesis of stationarity against a unit root alternative, when the series is subject to structural change at an unknown point in time. Three extant tests are reviewed which allow for an endogenously determined instantaneous structural break, and a related fourth procedure is introduced. We further propose tests which permit the structural change to be gradual rather than instantaneous, allowing the null hypothesis to be stationarity about a smooth transition in linear trend. The size and power properties of the tests are investigated, and the tests are applied to four economic time series.

I. Introduction

Hypothesis testing to determine whether economic time series are best modelled by (trend) stationary or unit root processes has been the subject of much recent research, and can be examined using two basic approaches. The first approach – 'unit root testing' – follows Dickey and Fuller (1979), and specifies the null hypothesis as a unit root process, testing against a stationary alternative. The second approach – 'stationarity testing' – reverses the null and alternative hypotheses, following tests first devised by Nyblom and

*The authors are very grateful to the Editor, Anindya Banerjee, and two anonymous referees for very helpful comments on earlier drafts of this paper.

JEL Classification numbers: C12, C22.

Mäkeläinen (1983) and MacNeill (1978), and subsequently generalized to allow the series to follow a general autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA) process by Kwiatkowski, Phillips, Schmidt and Shin (1992) – hereafter KPSS – and Leybourne and McCabe (1994, 1999).

Modifications to both sets of procedures are required when the series is subject to a structural break. In the unit root testing framework, Perron (1989, 1993) first developed procedures to conduct testing when a break occurs at a known point in time, while several authors have considered more general techniques which assume that the break date is unknown and is to be determined endogenously (see, e.g. Vogelsang and Perron, 1998; Harvey, Leybourne and Newbold, 2001). When the null hypothesis is that of stationarity with a structural break, Lee (1999), Lee and Strazicich (2001), Busetti and Harvey (2001, 2002) and Kurozumi (2002) propose KPSS-type tests which are modified to permit a break at either a known or an unknown date.

In practice, the implicit assumption that any structural change occurs instantaneously may be unrealistic. An alternative approach in the unit root testing context is proposed by Leybourne, Newbold and Vougas (1998), where the structural change is modelled by a smooth transition in linear trend. This procedure, which also allows the timing and speed of the transition to be determined endogenously, allows a series to gradually evolve from one trend regime to another, rather than undergo an instantaneous structural break.

In this paper, we focus on stationarity tests in the presence of endogenously determined structural change. We review three extant instant break tests, and introduce a related fourth procedure, deriving the test's asymptotic distribution under the null. Further, we propose two tests of the null hypothesis that the series is stationary about a smooth transition in linear trend, paralleling the Leybourne *et al.* (1998) unit root testing work. The tests are then compared by way of Monte Carlo simulations, investigating size and power when instantaneous breaks and smooth transitions are present. The structure of the paper is as follows: section II presents the instant break stationarity tests, while section III proposes the smooth transition test procedures; the simulation experiments are examined in section IV. In section V, the tests are applied to four economic time series: the *S&P 500* stock market index, UK GDP, British industrial production, and the dollar–sterling exchange rate. Section VI concludes.

II. Stationarity tests in the presence of breaks

Following Busetti and Harvey (2001), we consider four random walk models with different orders of deterministic components and structural breaks:

$$\text{Model 1} \quad y_t = \mu_t + \delta_\mu w_t' + \varepsilon_t \tag{1}$$

$$\text{Model 2} \quad y_t = \mu_t + \beta t + \delta_\mu w_t' + \delta_\beta (w_t' t) + \varepsilon_t \tag{2}$$

$$\text{Model 2a} \quad y_t = \mu_t + \beta t + \delta_\mu w_t' + \varepsilon_t \tag{3}$$

$$\text{Model 2b} \quad y_t = \mu_t + \beta t + \delta_\beta z_t' + \varepsilon_t \tag{4}$$

where $\mu_t = \mu_{t-1} + \eta_t$ with $\eta_t \sim \text{i.i.d.}(0, \sigma_\eta^2)$, $\varepsilon_t \sim \text{i.i.d.}(0, \sigma^2)$, $w_t{}' = 1(t > \tau')$ and $z_t{}' = 1(t > \tau')(t - \tau')$, with $1(\cdot)$ being the indicator function and τ' the true break point. Model 1 has no trend and a break in level; models 2a, 2b and 2 contain trends and are subject to, breaks in level only, slope only, and both level and slope, respectively. The models can be generalized to allow for autocorrelation by permitting ε_t to follow a more general stationary process, but for purposes of tractability in our analysis, we shall assume that this is unnecessary.

Busetti and Harvey (2001), and Lee and Strazicich (2001) for models 1 and 2, propose tests of the null hypothesis of stationarity $(H_0 : \sigma_\eta^2 = 0)$ against a unit root alternative $(H_1 : \sigma_\eta^2 > 0)$ when the true break date τ' is unknown. If the true break date was *known*, the appropriate test statistic (the locally best invariant test for Gaussian η_t and ε_t) would be a KPSS-type statistic constructed from the residuals of the relevant regression among equations (1)–(4) above, with μ_t replaced by a constant μ:

$$\xi_i(\lambda') = \frac{\sum_{t=1}^{T} \left(\sum_{s=1}^{t} e_s \right)^2}{T^2 \hat{\sigma}^2} \quad i = 1, 2, 2\text{a}, 2\text{b} \tag{5}$$

where e_t denotes the residuals, $\lambda' = \tau'/T$, and $\hat{\sigma}^2$ is the usual least squares estimator of σ^2. Now, given that the timing of the break is *unknown*, the statistic (5) can be calculated for all possible break dates τ and the break point selected as that giving the most favourable result for the null of (trend) stationarity about a structural break (cf. Zivot and Andrews, 1992, in the unit root testing context). This yields the following test statistic:

$$\tilde{\xi}_i = \inf_{\lambda \in \Lambda} \xi_i(\lambda) \quad i = 1, 2, 2\text{a}, 2\text{b} \tag{6}$$

where $\lambda = \tau/T$ and Λ is a closed subset of the interval (0,1). The break date is implicitly estimated by $\hat{\tau} = \hat{\lambda} T$, with $\hat{\lambda}$ obtained from:

$$\hat{\lambda} = \arg\min_{\lambda \in \Lambda} \, \xi_i(\lambda) \quad i = 1, 2, 2\text{a}, 2\text{b}. \tag{7}$$

The authors differ in their approach to obtaining critical values for these tests. Busetti and Harvey (2001) propose an *unconditional* procedure with regard to the break date: limit distributions are derived under the null for the statistics (6) by making use of a 'shift assumption' that the magnitudes of the breaks δ_μ and δ_β decrease to zero with the sample size at rates faster than $T^{-1/2}$ and $T^{-3/2}$, respectively. This assumption ensures that the distributions do not depend on either the break fraction or other nuisance parameters, and, of course, encompasses the case where no break actually occurs. Clearly, the value of this approach is highly dependent on the viability of the shift assumption in practice – if violated, dependencies on the break fraction and break magnitude will be introduced into the asymptotic distributions.[1] Lee and Strazicich (2001), however, suppose that the true break fraction λ' is consistently estimated by equation (7), and propose an approach that is *conditional* on the break timing, using critical values that correspond to tests with a break at a *known* point in time. Critical values – which depend on the value of λ' – are thus drawn from the limiting distributions of $\xi_i(\lambda')$ in equation (5).[2] The merit of this procedure rests primarily on the performance of equation (7) in estimating the true break fraction, and also on the implicit assumption that a break does actually occur.

In addition to the selection rule (7), other break date estimation procedures can be considered. Lee (1999) suggests use of the Schwarz Bayesian criterion (SBC) in the wider context of potential multiple breaks, and uses the criterion to simultaneously determine the *number* of breaks as well as their timings. Kurozumi (2002) and Busetti and Harvey (2002) propose estimating the relevant regression for all possible break dates, and using a break date which minimizes the residual sum of squares; this procedure will clearly give identical estimates to Lee's SBC approach when only one break is considered. In the parallel context of testing the null of a unit root in the presence of a break at an unknown time, selection procedures based on the significance of coefficients on the relevant break dummy variables have also been used (see, e.g. Vogelsang and Perron, 1998); indeed, Harvey *et al.* (2001) find a modified version of this method to be superior to the Zivot and Andrews (1992) minimum test statistic approach and the use of SBC in important cases. In the present context of testing a stationary null, break date selection based on the significance of break dummy variable coefficients translates to the following criteria, following estimation of the relevant regression for all possible break dates:

[1]The null limiting distributions for this unconditional procedure, along with the corresponding critical values, are given in Busetti and Harvey (2001), although the result for model 2a contains an error – the corrected distribution and critical values for this case, as well as finite sample critical values for all four models, are provided by Harvey and Mills (2002).

[2]These are reported for models 1 and 2 in Lee and Strazicich (2001) and for all four models in Busetti and Harvey (2001).

$$\begin{aligned} \hat{\lambda} &= \arg\max_{\lambda\in\Lambda} |t_{\hat{\delta}_\mu}(\lambda)| && \text{for models 1, 2a} \\ \hat{\lambda} &= \arg\max_{\lambda\in\Lambda} F_{\hat{\delta}_\mu,\hat{\delta}_\beta}(\lambda) && \text{for model 2} \\ \hat{\lambda} &= \arg\max_{\lambda\in\Lambda} |t_{\hat{\delta}_\beta}(\lambda)| && \text{for model 2b} \end{aligned} \tag{8}$$

where $t_{\hat{\delta}_\mu}(\lambda)$ and $t_{\hat{\delta}_\beta}(\lambda)$ are the *t*-ratios on δ_μ and δ_β, respectively, and $F_{\hat{\delta}_\mu,\hat{\delta}_\beta}(\lambda)$ is the *F*-statistic for testing the joint significance of δ_μ and δ_β, all for a given break date $\tau = \lambda T$. Of course, unlike in the unit root testing framework, use of equation (8) will always give identical break date estimates to the minimum residual sum of squares and SBC criteria discussed above. The stationarity test statistics which follow from this approach are simply obtained by computing $\xi_i(\hat{\lambda})$ using equation (5), where $\hat{\lambda}$ is calculated according to equation (8).

As with the minimum stationarity test statistic procedure (6), there are two possible approaches for obtaining critical values for the $\xi_i(\hat{\lambda})$ test statistics. One approach, which we develop here, is to derive the null limit distributions of the statistics under the shift assumption of Busetti and Harvey (2001), thereby providing an *unconditional* procedure with regard to the break date. Using version (8) of the equivalent estimation criteria discussed above to maintain consistency with the preferred approach in the unit root testing context, the null asymptotic distributions are (see Appendix for proof):

$$\xi_i(\hat{\lambda}) \Rightarrow \int_0^1 [B_i(r,\hat{\lambda})]^2 \mathrm{d}r \quad i = 1, 2, 2\mathrm{a}, 2\mathrm{b} \tag{9}$$

where the $B_i(r,\hat{\lambda})$ are as defined in Busetti and Harvey (2001) for $i = 1, 2, 2\mathrm{b}$, and Harvey and Mills (2002) for $i = 2\mathrm{a}$, except with λ replaced by $\hat{\lambda}$, where

$$\hat{\lambda} = \arg\max_{\lambda\in\Lambda} |Z_i(\lambda)|$$

with

$$Z_1(\lambda) = \sqrt{\tfrac{\lambda}{1-\lambda}}[W(1) - \tfrac{1}{\lambda}W(\lambda)]$$

$$\begin{aligned} Z_2(\lambda) = \tfrac{2}{\lambda^3(1-\lambda)^5}\ &[(1-3\lambda+3\lambda^2)H_1^2 + 3(1-3\lambda+4\lambda^2)^2 H_1 H_2 \\ &+ 3(1-2\lambda+4\lambda^2)H_2^2] \end{aligned}$$

$$Z_{2\mathrm{a}}(\lambda) = \sqrt{\tfrac{\lambda}{(1-\lambda)(1-3\lambda+3\lambda^2)}}\left\{W(1) - \tfrac{1}{\lambda}W(\lambda) - 6(1-\lambda)\left[\int_0^1 r\,\mathrm{d}W(r) - \tfrac{1}{2}W(1)\right]\right\}$$

TABLE 1

Critical values of $\xi_i(\hat{\lambda})$ tests for stationarity with an endogenously determined structural break

	Model 1			Model 2		
	10%	5%	1%	10%	5%	1%
$T = 50$	0.167	0.209	0.311	0.047	0.057	0.078
$T = 100$	0.170	0.212	0.320	0.048	0.057	0.081
$T = 200$	0.177	0.219	0.323	0.048	0.057	0.081
$T = \infty$	0.177	0.221	0.326	0.049	0.057	0.081
	Model 2a			Model 2b		
	10%	5%	1%	10%	5%	1%
$T = 50$	0.070	0.085	0.122	0.062	0.075	0.108
$T = 100$	0.071	0.086	0.119	0.062	0.076	0.112
$T = 200$	0.072	0.086	0.122	0.062	0.077	0.107
$T = \infty$	0.073	0.088	0.122	0.062	0.075	0.106

Notes: Model 1 has no trend and allows for a break in level; models 2a, 2b and 2 contain trends and allow for breaks in level only, trend only, and level and trend, respectively [see equations (1)–(4)]. $\xi_i(\hat{\lambda})$ denotes tests for stationarity allowing for an instantaneous break, where the break date is estimated according to the significance of break dummy variable coefficients. The reported critical values relate to the unconditional procedure.

$$Z_{2b}(\lambda) = \sqrt{\tfrac{3}{\lambda^3(1-\lambda)^3}}\Big\{\lambda^2[(\lambda-1)W(1) - [W(1) - \tfrac{1}{\lambda}W(\lambda)]] \\ -(1-\lambda)^2(1+2\lambda)\int_0^\lambda r\,\mathrm{d}W(r) + \lambda^2(3-2\lambda)\int_\lambda^1 r\,\mathrm{d}W(r)\Big\}$$

and

$$H_1 = 2\lambda^2(1+\lambda+\lambda^2)W(1) - 2\lambda(1-2\lambda+4\lambda^2)W(\lambda) \\ + 3(1-\lambda)^3\int_0^\lambda r\,\mathrm{d}W(r) - 3\lambda^2(1+\lambda)\int_\lambda^1 r\,\mathrm{d}W(r)$$

$$H_2 = -\lambda^3(1+\lambda)W(1) + \lambda(1-3\lambda+4\lambda^2)W(\lambda) \\ - 2(1-\lambda)^3\int_0^\lambda r\,\mathrm{d}W(r) + 2\lambda^3\int_\lambda^1 r\,\mathrm{d}W(r)$$

where $W(r)$ is a standard Wiener process. Asymptotic and finite sample critical values for these tests were obtained by Monte Carlo simulation and are given in Table 1. The asymptotic critical values were generated by direct simulation of the limiting functionals in equation (9) using discrete approximations for $T = 500$. The space of values for λ is restricted to a closed subset of (0, 1); in our simulations here and in section IV, we restricted

λ to lie between the conventionally chosen points (0.2, 0.8). The finite sample critical values were obtained by repeated application of the test $\xi_i(\hat{\lambda})$, using equation (8) to estimate the break date, to a stationary generating process [$y_t = \varepsilon_t \sim$ n.i.d.(0, 1)].[3] Here and throughout the paper, simulations were programmed in GAUSS and conducted using 10,000 replications. With the exception of model 1, the finite sample critical values are very close to those in the limit. As with the Busetti and Harvey (2001) unconditional test, the potential limit of this new approach is the credibility of the shift assumption in realistic scenarios, as will be considered later in the paper.

The second approach, proposed by Lee (1999), Kurozumi (2002) and Busetti and Harvey (2002), is to adopt a procedure that is *conditional* on the break timing, using critical values associated with tests involving a break at a known date, given the fact that equation (8)/minimum residual sum of squares/SBC will estimate the true break fraction λ' superconsistently. As with the Lee and Strazicich (2001) unconditional test using equation (7), the performance of this method depends on how well λ' is estimated, and behaviour when a break is not actually present in the series.

III. Stationarity tests in the presence of smooth transitions

The stationarity tests of the previous section all assume that when there is a break, the change occurs instantaneously. However, as argued by Leybourne *et al.* (1998) in the context of testing a unit root null hypothesis, this may be unrealistic in many economic applications. A more appealing approach would be to allow any structural change to occur gradually, with a smooth, rather than instantaneous, transition between two deterministic regimes. Leybourne *et al.* (1998) propose unit root tests where the alternative hypothesis is stationary about such a smooth transition in linear trend. These authors make use of the logisitic smooth transition function to model the change, following the work of Bacon and Watts (1971), Granger and Teräsvirta (1993) and Lin and Teräsvirta (1994).

Applying these concepts to the stationarity testing context, we can consider variants of the models (1)–(3), where the structural break terms are replaced by smooth transitions:[4]

[3]This method of obtaining finite sample critical values (repeated application of the test to a stationary series without breaks) parallels the situation in the limit where, under the shift assumption, there is no dependence on break parameters. This same approach is used by Harvey and Mills (2002) to obtain finite sample critical values for the Busetti and Harvey (2001) tests.

[4]Smooth transition models do not apply well to the piecewise linear trend model 2b, since this model relies on a single point at which the two regimes meet; a smooth transition variant of this type of process is better modelled by a transition in intercept and trend, i.e. model 2.

$$\text{Model 1} \quad y_t = \mu_t + \delta_\mu S'_t + \varepsilon_t \tag{10}$$

$$\text{Model 2} \quad y_t = \mu_t + \beta t + \delta_\mu S'_t + \delta_\beta (S'_t t) + \varepsilon_t \tag{11}$$

$$\text{Model 2a} \quad y_t = \mu_t + \beta t + \delta_\mu S'_t + \varepsilon_t \tag{12}$$

where μ_t and ε_t are as defined below equations (1)–(4), and $S_t{}'$ is the logistic smooth transition function:

$$S'_t = [1 + \exp\{-\gamma'(t - \lambda' T)/\hat{\sigma}(t)\}]^{-1} \tag{13}$$

with $\sigma(t)$ denoting the standard deviation of the transition variable, i.e. the time trend.

The midpoint of the transition is given by $\lambda' T$, and the transition speed is controlled by the parameter γ'. At one extreme, $\gamma' = 0$ implies that no transition occurs at all, while in the limit $\gamma' \to \infty$, the transition occurs instantaneously: thus no structural change and an instantaneous structural break are special cases of the above models. The term $\hat{\sigma}(t)$ in equation (13) is introduced, following Granger and Teräsvirta (1993, p. 123), to provide an $O(T)$ scaling to γ', allowing transition speeds for different sample sizes to be interpreted using γ' independent of T. The transitions considered here are *symmetric* about $\lambda' T$; further generalization to admit *asymmetric* smooth transitions can be performed using the generalized logistic function, as examined by Sollis, Leybourne and Newbold (1999) for testing a unit root null. Such extensions, however, go beyond the scope of this paper.

Tests of the null hypothesis of stationarity, where the transition speed and timing are determined endogenously, can be conducted using a two-stage procedure. In the first stage, the appropriate regression model based on equations (10)–(12) is estimated by nonlinear least squares (NLS), with μ_t replaced by a constant μ. The residual sum of squares function can be minimized analytically over the regression parameters (μ, β, δ_μ, δ_β) and numerically over the transition parameters (γ, λ). In the second stage, a KPSS-type test statistic is computed:

$$s_i = \frac{\sum_{t=1}^{T} \left(\sum_{s=1}^{t} e_s\right)^2}{T^2 \hat{\sigma}^2} \quad i = 1, 2, 2\text{a} \tag{14}$$

where e_t denotes the residuals from the first-stage regression. Conceptually, this method is most similar to the instant break stationarity tests where the break date is selected using equation (8), as in both cases the most likely timing of the structural change is explicitly determined as a first step, with the stationarity tests subsequently computed conditional on this choice. As with the instant break tests, the test statistics can be modified to allow for additional autocorrelation in the series, but here we shall assume that such augmentation is unnecessary.

Critical values for these tests can be obtained in two ways, paralleling the unconditional and conditional approaches adopted in the case of instantaneous

TABLE 2

Critical values of s_i tests for stationarity with an endogenously determined smooth transition in linear trend

	Model 1			*Model 2*			*Model 2a*		
	10%	*5%*	*1%*	*10%*	*5%*	*1%*	*10%*	*5%*	*1%*
$T = 50$	0.180	0.230	0.354	0.051	0.062	0.091	0.063	0.076	0.107
$T = 100$	0.175	0.223	0.348	0.050	0.061	0.084	0.063	0.077	0.106
$T = 200$	0.175	0.223	0.354	0.048	0.059	0.086	0.063	0.075	0.107
$T = 500$	0.183	0.235	0.379	0.049	0.061	0.084	0.064	0.076	0.106

Notes: Model 1 has no trend and allows for a smooth transition in level; models 2a and 2 contain trends and allow for smooth transitions in level only, and level and trend, respectively [see equations (10)–(12)]. s_i denotes tests for stationarity allowing for a smooth transition. The reported critical values relate to the unconditional procedure.

breaks. The difference, however, is that analytical derivations of the asymptotic distributions of the tests are not possible using standard techniques, because of the lack of closed-form expressions for the NLS estimators, and the need for knowledge of the limiting behaviour of terms such as $\sum_{t=1}^{T} S_t' y_t$. The unconditional approach in the instantaneous break framework rests on the Busetti and Harvey (2001) shift assumption. Although asymptotic distributions are not directly evaluated in the smooth transition case, it is to be expected that the shift assumption will be equally important in preventing the null limiting distributions of the smooth transition tests from being dependent on nuisance parameters. We implicitly make this assumption, therefore, for the unconditional approach, and derive finite sample critical values in the same way as for the instantaneous break tests of this type, i.e. using Monte Carlo simulation, repeatedly applying the tests of equation (14) to a stationary generating process [$y_t = \varepsilon_t \sim$ n.i.d.(0, 1)]. Finite sample critical values generated in this way are provided in Table 2, and results for $T = 500$ can be taken as approximate asymptotic critical values for the tests.[5] This unconditional procedure clearly nests the case where no transitions are present in the series.

If a smooth transition does occur, the NLS estimation will provide consistent estimates of the parameters, motivating the possibility of a conditional approach regarding critical values for the tests. This second method reasons that the first stage optimization will provide consistent estimates of γ' and λ', in which case critical values corresponding to tests with a smooth transition with *known* speed and midpoint might be employed. As with the conditional instantaneous

[5]For each replication (here and in the section IV simulations), starting values for the nonlinear estimation were obtained from a grid-search procedure, with values for γ and λ given by $\gamma \in \{1.5, 3, 7.5, 15, 30\}$ and $\lambda \in \{0.1, 0.2, \ldots, 0.9\}$ respectively. Numerical optimization was performed using the Broyden, Fletcher, Goldfarb and Shanno algorithm in the OPTMUM subroutine for GAUSS; in any cases where the algorithm produced estimates of γ and λ which resulted in the transition function S_t being constant for the whole sample, the Moore-Penrose generalized inverse was employed.

TABLE 3

Critical values of s_i tests for stationarity with a known smooth transition in linear trend: 5% level, $T = 500$

λ'	$\gamma' = 1.5$	$\gamma' = 3$	$\gamma' = 7.5$	$\gamma' = 15$	$\gamma' = 30$	$\gamma' = 60$
(A) Model 1						
0.01	0.196	0.278	0.374	0.412	0.426	0.436
0.10	0.185	0.253	0.324	0.349	0.358	0.364
0.20	0.171	0.219	0.270	0.286	0.295	0.299
0.30	0.158	0.184	0.216	0.228	0.235	0.238
0.40	0.148	0.158	0.178	0.188	0.192	0.196
0.50	0.144	0.148	0.163	0.174	0.179	0.181
0.60	0.147	0.156	0.177	0.186	0.194	0.197
0.70	0.156	0.182	0.213	0.224	0.233	0.238
0.80	0.169	0.217	0.270	0.287	0.296	0.304
0.90	0.183	0.252	0.329	0.357	0.368	0.372
0.99	0.193	0.279	0.377	0.415	0.433	0.440
(B) Model 2						
0.01	0.061	0.074	0.106	0.126	0.138	0.142
0.10	0.060	0.070	0.098	0.110	0.117	0.120
0.20	0.058	0.065	0.083	0.091	0.094	0.096
0.30	0.057	0.059	0.069	0.073	0.076	0.077
0.40	0.056	0.055	0.058	0.062	0.063	0.065
0.50	0.056	0.053	0.055	0.058	0.059	0.061
0.60	0.056	0.055	0.059	0.062	0.064	0.065
0.70	0.057	0.059	0.069	0.075	0.077	0.078
0.80	0.058	0.065	0.082	0.089	0.093	0.095
0.90	0.059	0.070	0.095	0.110	0.117	0.118
0.99	0.060	0.074	0.106	0.124	0.134	0.140
(C) Model 2a						
0.01	0.087	0.099	0.122	0.135	0.140	0.142
0.10	0.085	0.092	0.109	0.116	0.119	0.121
0.20	0.084	0.085	0.093	0.098	0.100	0.101
0.30	0.086	0.084	0.090	0.095	0.099	0.101
0.40	0.101	0.103	0.112	0.116	0.119	0.120
0.50	0.130	0.129	0.129	0.130	0.132	0.133
0.60	0.103	0.105	0.113	0.118	0.121	0.122
0.70	0.088	0.085	0.091	0.096	0.099	0.102
0.80	0.084	0.085	0.092	0.095	0.098	0.100
0.90	0.085	0.092	0.107	0.114	0.118	0.119
0.99	0.087	0.098	0.122	0.133	0.140	0.141

Notes: See notes to Table 2 for model and test definitions. λ' and γ' denote the smooth transition midpoint fraction and speed, respectively. The reported critical values relate to the conditional procedure.

break procedures, the resulting test's performance will be dependent on the performance of the transition parameter estimators, as well as behaviour when no transition actually occurs. The critical values will of course depend on both γ' and λ', and can be generated by simulation. Table 3 contains a grid of such values at the 5% level for a sample size of $T = 500$; these were generated by repeated application of equation (14) to a generating model of stationarity about a smooth transition in linear trend, i.e. equations (10)–(12) with $\sigma_\eta^2 = 0$, treating the transition parameters as known in the first-stage regressions. The Table 3 values approximate the asymptotic critical values for the conditional procedure; in unreported experiments, we also found critical values for $T = 100$ to differ little from those in Table 3, especially for models 2 and 2a.[6]

IV. Monte Carlo simulations

In order to examine the behaviour of the tests outlined in sections II and III, and compare their relative performance, we conducted a number of Monte Carlo experiments. Let $\tilde{\xi}_i$ and $\xi_i(\hat{\lambda})$ denote tests employing break date selection rules (7) and (8), respectively, with critical values consistent with the unconditional approach described in section II, i.e. using critical values from Harvey and Mills (2002) and Table 1, respectively. Similarly, let s_i denote the unconditional smooth transition stationarity test, using the critical values of Table 2. Then, let $\tilde{\xi}_i^*$, $\xi_i(\hat{\lambda})^*$ and s_i^* denote tests using the same test statistics as above, but with critical values corresponding to conditional tests where the timing (and speed for smooth transitions) of structural change is assumed known, i.e. using critical values similar to those of Table 3 for the smooth transition tests (critical values for the exact sample size involved were actually used, with a slightly more detailed grid), and similarly simulated critical values for the instant break tests for a range of λ' values; the precise critical values used in a given replication were then obtained by linear interpolation.

The generating processes used for the size and power simulations are given by equations (1)–(4) for series containing instantaneous breaks[7] and equation (12) for a representative case involving smooth transitions, with $\sigma_\eta^2 = 0$ for size experiments and $\sigma_\eta^2 = 0.01, 0.1, 1$ for power experiments. We set $\mu_0 = \beta = 0$ and $\varepsilon_t \sim$ n.i.d.(0, 1) throughout, and unless otherwise stated, experiments were conducted using a typical sample size of $T = 100$. Simulations were conducted for breaks/transitions occurring early, late and in

[6]If the null hypothesis of stationarity about a smooth transition is not rejected, a related issue is to subsequently test whether a transition is actually present in the series, i.e. testing the null of linearity ($\gamma' = 0$) under an assumption of stationarity. Granger and Teräsvirta (1993, Chapter 6) provide a detailed discussion of such tests.

[7]An exception is that for model 2, $w_t{}'t$ in equation (2) is replaced by $z_t{}'$ so as to separate out the impacts of the level and trend components of the breaks.

the middle of the sample, but as the results were found to be approximately symmetric about $\lambda' = 0.5$, we report $\lambda' = 0.25$, 0.5 only. All six tests are reported in each case, but note that as we do not consider a model 2b version of the smooth transition tests, we applied the s_2 and s_2^* tests when the generating process was equation (3). Where they are comparable, our results and comments concur with those of Busetti and Harvey (2002) and Lee and Strazicich (2001).

Size and break date estimation when an instantaneous break occurs

Table 4 provides results for size experiments when a break occurs under the null hypothesis. Examining first the unconditional instant break tests, although (by construction) both $\tilde{\xi}_i$ and $\xi_i(\hat{\lambda})$ are correctly sized when no break occurs, the tests suffer from size distortions when breaks are present under the null. The distortions are mostly those of oversize and vary according to the true break fractions. These distortions and dependencies on λ' clearly highlight that the shift assumption underpinning the unconditional tests is inappropriate for the break magnitudes considered here. It appears that the shift assumption would only be viable for much smaller breaks, but as the magnitudes considered in Table 4 are typical of those arising in economic time series, the results question the validity of the unconditional tests in practical applications. The conditional test $\tilde{\xi}_i^*$, which uses break date estimator (7), also suffers from size distortions. This procedure is always undersized, often to a quite severe degree. In order to examine whether the size distortions of these three procedures are small sample effects, we also conducted experiments for larger sample sizes. The distortions were found to persist in larger samples (with the exception of $\tilde{\xi}_{2b}^*$), increasing the uncertainty as to the reliability of these tests in practice.

In contrast, the conditional procedure using equation (8) to estimate the break date, $\xi_i(\hat{\lambda})^*$, performs extremely well for series containing breaks. Approximate correct size is almost always achieved for the models, break magnitudes and break timings considered. Moreover, unreported simulations show that any deviations from nominal size disappear as the sample size grows, yielding a well-behaved test with very appealing size properties.

The only drawback concerns the performance of the test when no break is present under the null – in these circumstances the test is undersized. In fact, simulations show that, asymptotically, the estimator (8) places the break at the beginning or end of the series when no break occurs. This feature ensures that the test can still be theoretically valid in the case of no break, as the critical values for break fractions close to zero and one converge to those of the standard no-break KPSS test. However, two caveats should be highlighted: first, the degree of trimming selected for the space of values for λ (in our simulations, 20%) constrains the fraction estimates, so that unless close to zero trimming is used, the endpoints cannot practically be selected by equation (8);

TABLE 4

Empirical sizes of nominal 5% level stationarity tests when a structural break occurs under the null: T = 100

δ_β	δ_μ	λ'	$\tilde{\xi}_1$	$\tilde{\xi}_1^*$	$\xi_1(\hat{\lambda})$	$\xi_1(\hat{\lambda})^*$	s_1	s_1^*
(A) Model 1								
—	0.0	—	5.08	0.07	5.01	3.70	5.02	2.72
	2.5	0.25	17.87	1.46	9.21	4.73	6.87	3.92
		0.50	10.28	1.31	2.71	4.84	1.65	4.17
	5.0	0.25	17.62	1.30	9.72	4.94	7.84	4.79
		0.50	10.46	1.48	2.77	4.98	2.02	4.78
	10.0	0.25	14.81	0.37	9.78	4.99	8.05	4.48
		0.50	9.92	1.34	2.76	4.96	2.11	4.82

δ_β	δ_μ	λ'	$\tilde{\xi}_{2a}$	$\tilde{\xi}_{2a}^*$	$\xi_{2a}(\hat{\lambda})$	$\xi_{2a}(\hat{\lambda})^*$	s_{2a}	s_{2a}^*
(B) Model 2a								
—	0.0	—	5.08	0.03	4.81	1.51	5.03	0.71
	2.5	0.25	21.97	1.79	8.83	4.73	8.71	2.93
		0.50	23.83	0.58	16.34	4.86	17.63	3.77
	5.0	0.25	22.42	1.85	9.01	4.96	11.83	4.51
		0.50	22.29	0.25	16.78	5.02	20.96	4.86
	10.0	0.25	23.02	1.61	9.05	4.98	12.70	4.73
		0.50	22.83	0.02	16.82	5.01	20.49	4.93

δ_β	δ_μ	λ'	$\tilde{\xi}_{2b}$	$\tilde{\xi}_{2b}^*$	$\xi_{2b}(\hat{\lambda})$	$\xi_{2b}(\hat{\lambda})^*$	s_2	s_2^*
(C) Model 2b								
0	—	—	5.14	0.58	5.00	1.55	4.92	0.67
1	—	0.25	7.67	0.90	9.90	3.43	8.80	2.78
		0.50	2.52	0.49	3.25	2.07	2.84	3.34
2	—	0.25	11.98	1.41	12.51	4.63	13.54	4.68
		0.50	9.64	2.69	6.61	4.67	3.74	4.46
4	—	0.25	19.07	4.99	13.21	5.12	17.02	6.00
		0.50	11.57	4.98	6.93	4.98	3.81	4.16

δ_β	δ_μ	λ'	$\tilde{\xi}_2$	$\tilde{\xi}_2^*$	$\xi_2(\hat{\lambda})$	$\xi_2(\hat{\lambda})^*$	s_2	s_2^*
(D) Model 2								
0	0.0	—	4.67	0.03	5.07	1.23	4.92	0.67
1	2.5	0.25	31.77	1.91	18.86	4.93	12.94	4.10
		0.50	14.20	0.87	6.36	4.97	3.12	3.43
	5.0	0.25	30.49	1.23	18.63	4.91	15.34	4.78
		0.50	15.66	1.17	6.47	5.06	3.96	4.35
	10.0	0.25	27.93	0.28	18.62	4.93	15.66	5.17
		0.50	18.86	1.65	6.54	5.11	4.33	4.74
2	2.5	0.25	28.70	0.94	18.70	4.92	13.49	4.12
		0.50	13.56	0.87	6.33	5.05	3.58	4.03
	5.0	0.25	26.83	0.83	18.66	4.91	14.87	4.93
		0.50	11.02	0.26	6.49	5.08	3.93	4.28
	10.0	0.25	26.69	0.15	18.62	4.93	15.58	5.17
		0.50	19.00	1.31	6.54	5.11	4.24	4.65

TABLE 4

(continued)

δ_β	δ_μ	λ'	$\tilde{\xi}_2$	$\tilde{\xi}_2^*$	$\xi_2(\hat{\lambda})$	$\xi_2(\hat{\lambda})^*$	s_2	s_2^*
4	2.5	0.25	26.45	0.59	18.83	4.98	14.70	4.60
		0.50	12.54	0.80	6.40	4.97	3.57	3.94
	5.0	0.25	18.07	0.18	18.65	4.91	14.84	4.73
		0.50	6.82	0.11	6.46	5.06	4.08	4.48
	10.0	0.25	18.08	0.05	18.62	4.93	15.80	4.89
		0.50	7.49	0.00	6.54	4.11	4.23	4.51

Notes: See notes to Tables 1 and 2 for model and test definitions. Tests denoted with and without an '*' use critical values corresponding to the conditional and unconditional procedures, respectively. λ' denotes the break fraction, while δ_μ and δ_β denote the magnitudes of the breaks in level and trend, respectively.

secondly, convergence of $\hat{\lambda}$ to the endpoints is an asymptotic property, and occurs only to a limited extent in finite samples. These problems lead to the observed undersize for $\xi_i(\hat{\lambda})^*$ in the tables. Despite this drawback, our primary concern is behaviour when breaks *do* occur; if it is known that a break is not present in a series, tests explicitly designed for such situations (e.g. KPSS) should be used instead. On the other hand, if $\xi_i(\hat{\lambda})^*$ is used in practice and the estimated break fraction is found to be close to either zero or one (or more accurately the extremes of the space of values considered for λ), this would motivate further exploration as to whether a break is actually present in the series. [Note that the above arguments do not apply to $\tilde{\xi}_i^*$, as the estimator (7) does not converge to the endpoints in the no-break case.]

The results of Table 4 also show that tests based on equation (8) are almost always better sized than those based on equation (7). This arises as equation (8) is a substantially superior break fraction estimator, and is of particular importance to the conditional tests because of their explicit reliance on a good estimate of the break timing. Table 5 reports, for representative cases, the sample means and variances of estimated break fractions from the size simulations, in addition to results for the same experiments using larger samples of $T = 200$ and $T = 400$. The mean is always more accurately estimated by equation (8) than (7), and the variance is always substantially lower, yielding a more efficient estimator. Moreover, in line with theory, evidence exists that equation (8) is a superconsistent estimator, converging to λ' at rate T (or faster), with the variance reducing by at least a factor of 4 as the sample size doubles. The inferior estimator (7) is still consistent for λ', but converges at a slower rate, with evidence of superconsistency only present for model 2b. These findings explain the difference in size performance for the two conditional tests. For models 1, 2a and 2, equation (7) does not achieve rate T consistency, and thus does not converge to the true value fast enough to prevent test size distortions. Instead, given that the nature of the estimator (7) is to choose the minimum

TABLE 5

Means and variances of estimated break fractions using estimation criteria (7) and (8) when a structural break occurs under the null

			$\lambda' = 0.25$		$\lambda' = 0.50$	
δ_β	δ_μ	T	(7)	(8)	(7)	(8)
(A) Model 1						
—	2.5	100	0.2514 (4.37×10^{-4})	0.2503 (6.62×10^{-5})	0.5000 (2.92×10^{-4})	0.5000 (6.64×10^{-5})
		200	0.2503 (1.68×10^{-4})	0.2500 (1.58×10^{-5})	0.5001 (1.18×10^{-4})	0.5000 (1.54×10^{-5})
		400	0.2495 (6.54×10^{-5})	0.2500 (3.85×10^{-6})	0.4999 (5.18×10^{-5})	0.5000 (3.73×10^{-6})
(B) Model 2a						
—	2.5	100	0.2508 (9.94×10^{-4})	0.2506 (2.84×10^{-4})	0.4998 (7.44×10^{-4})	0.4999 (1.08×10^{-4})
		200	0.2496 (8.81×10^{-5})	0.2499 (1.74×10^{-5})	0.5001 (1.52×10^{-4})	0.5000 (1.54×10^{-5})
		400	0.2496 (3.68×10^{-5})	0.2499 (3.99×10^{-6})	0.5000 (6.29×10^{-5})	0.5000 (3.72×10^{-6})
(C) Model 2b						
1	—	100	0.2502 (5.58×10^{-5})	0.2501 (2.92×10^{-5})	0.5000 (2.97×10^{-5})	0.5000 (2.08×10^{-5})
		200	0.2500 (6.96×10^{-6})	0.2500 (3.09×10^{-6})	0.5000 (3.46×10^{-6})	0.5000 (1.96×10^{-6})
		400	0.2500 (8.04×10^{-7})	0.2500 (2.00×10^{-7})	0.5000 (2.12×10^{-7})	0.5000 (7.13×10^{-8})
(D) Model 2						
1	2.5	100	0.2374 (6.54×10^{-4})	0.2470 (1.21×10^{-4})	0.4771 (7.40×10^{-4})	0.4964 (1.42×10^{-4})
		200	0.2300 (2.16×10^{-4})	0.2484 (3.31×10^{-5})	0.4846 (2.34×10^{-4})	0.4983 (3.28×10^{-5})
		400	0.2381 (5.16×10^{-5})	0.2492 (7.99×10^{-6})	0.4911 (7.00×10^{-5})	0.4992 (7.56×10^{-6})

Notes: Main cell entries are means; estimated variances are given in parentheses. See notes to Tables 1 and 4 for model and parameter definitions. Criteria (7) and (8) estimate the break date according to minimization of the stationarity test statistic, and the significance of break dummy variable coefficients, respectively.

possible test statistic, a tendency exists for $\tilde{\xi}_i^*$ to underreject the null hypothesis. In combination with the size results of Table 4, it is clear that $\xi_i(\hat{\lambda})^*$ has substantial advantages over the other instant break tests.

Returning now to Table 4, we can also evaluate the size behaviour of the smooth transition tests. The unconditional s_i test is subject to size distortions which generally worsen with the magnitude of the break; these distortions persist for all cases except where no break occurs, in which case the test is correctly sized by construction. [As might be expected, the pattern of oversize and undersize is very similar to that for the related $\xi_i(\hat{\lambda})$ test.] The results strongly suggest that the critical values used for the s_i test (derived so as to be consistent with an implicitly assumed shift assumption) are inappropriate when breaks of interesting magnitude are present. However, the conditional test s_i^* has much better size performance, achieving approximate correct size in most cases. The test exhibits a little distortion for small breaks, but is only strongly undersized in the no break case.[8] Considering the additional

[8]When no structural change occurs, the conditional test would only remain valid if $\hat{\gamma} \to \infty$ and $\hat{\lambda}$ converged to the endpoints in the limit, since it is only under these circumstances that the test reduces to the standard no-break KPSS test. Unreported simulations show that although $\hat{\gamma}$ diverges, the estimated midpoint does not converge to the endpoints, making the test invalid in the case of no break/transition.

generality which this smooth transition test admits, the reliable size performance of s_i^* in the presence of *instant* breaks is encouraging and signals potential value for applications.

Power when an instantaneous break occurs

Table 6 presents, for model 1, estimated powers of the tests when an instant break occurs under the alternative hypothesis;[9] these results are also representative of models 2a, 2b and 2. Consider first results for the four instant break tests. When $\sigma_\eta^2 = 1$, there is little difference in power among the tests; however, for smaller values of σ_η^2, clearer rankings can be observed. The tests that generally have highest power are the unconditional procedures, with

TABLE 6

Estimated powers of nominal 5% level model 1 stationarity tests when a structural break occurs under the alternative: $T = 100$, $\delta_\mu = \theta\sqrt{(\sigma_\eta^2 + 2)/2}$

σ_η^2	θ	λ'	$\tilde{\xi}_1$	$\tilde{\xi}_1^*$	$\xi_1(\hat{\lambda})$	$\xi_1(\hat{\lambda})^*$	s_1	s_1^*
0.01	0.0	—	29.96	6.74	18.90	16.04	12.42	9.86
	2.5	0.25	63.74	29.70	54.04	45.43	37.07	31.89
		0.50	47.99	19.46	34.39	40.85	21.64	30.38
	5.0	0.25	61.67	24.33	54.77	46.21	50.67	43.88
		0.50	45.31	18.60	35.94	42.29	30.06	39.53
	10.0	0.25	54.75	15.30	54.81	46.28	51.95	45.10
		0.50	38.41	14.40	36.05	42.40	31.94	41.38
0.10	0.0	—	87.67	59.76	73.51	70.25	44.02	44.67
	2.5	0.25	93.65	71.33	85.59	81.07	57.38	57.75
		0.50	92.29	71.29	82.81	83.49	53.62	59.56
	5.0	0.25	94.33	73.12	91.25	86.83	73.64	73.20
		0.50	92.51	75.71	88.69	90.96	71.19	78.97
	10.0	0.25	91.65	62.26	91.51	87.15	89.52	86.60
		0.50	87.96	65.96	89.12	91.44	86.10	90.58
1.00	0.0	—	99.55	92.73	97.10	95.90	78.29	82.97
	2.5	0.25	99.50	92.86	97.37	96.04	79.12	83.97
		0.50	99.57	93.29	97.44	96.48	78.47	84.06
	5.0	0.25	99.71	93.35	98.17	96.92	81.58	86.56
		0.50	99.55	94.64	97.96	97.39	80.11	86.44
	10.0	0.25	99.58	90.79	99.24	98.47	87.90	91.30
		0.50	99.43	94.46	99.00	99.13	85.70	92.25

Notes: See notes to Tables 1, 2 and 4 for model, test and parameter definitions. σ_η^2 denotes the variance of the unit root component disturbances [see equations (1)–(4)].

[9]The same break sizes are considered as with the size simulations, although the magnitudes are scaled so as to represent 2.5, 5 and 10 times the standard deviation of the process each time. Moreover, the scaling is normalized so that if $\sigma_\eta^2 = 0$, the breaks would be identical to those used in the previous size experiments, i.e. $\delta_\mu = \theta\sqrt{(\sigma_\eta^2 + 2)/2}$ with $\theta = 2.5$, 5, 10.

$\tilde{\xi}_1$ outperforming $\xi_1(\hat{\lambda})$ when the break magnitude is moderate or zero, although this advantage disappears for larger break sizes. The problem with these procedures is that the high power comes at a price of size distortions when breaks occur under the null. Of the two conditional tests, $\xi_1(\hat{\lambda})^*$ has uniformly greater power than $\tilde{\xi}_1^*$, with the differences being quite dramatic. This is perhaps not surprising given the undersizing that $\tilde{\xi}_1^*$ exhibits when breaks occur under the null. Comparing $\xi_1(\hat{\lambda})^*$ with the unconditional tests, it is reassuring to see that the power losses associated with this test relative to $\tilde{\xi}_1$ and $\xi_1(\hat{\lambda})$ are not too severe. These losses become less marked the larger the break magnitude, and are less severe when a break occurs in the middle of the series; indeed, when $\lambda' = 0.5$, $\xi_1(\hat{\lambda})^*$ outperforms $\xi_1(\hat{\lambda})$ for all non-zero breaks, and $\tilde{\xi}_1$ for large breaks. Thus the cost in terms of power of employing a test with reliable size is relatively small in most circumstances, adding to the case for using $\xi_i(\hat{\lambda})^*$ in practice.

With regard to the smooth transition tests, there is little to separate s_1 and s_1^*: neither test has uniformly greater power when $\sigma_\eta^2 = 0.01, 0.1$, although s_1^* does have some power advantage when $\sigma_\eta^2 = 1$. Given the size problems associated with s_1 when a break occurs under the null, these results confirm the value of the conditional s_1^* version of the test when a break is actually present. Both tests generally have lower power than the instant break tests and the differences are more marked for small break magnitudes than for large. Some power loss is expected due to the additional generality that the smooth transition tests provide, but given the potential advantages of this testing approach, the power losses are not too severe, particularly when compared to the reliably sized $\xi_1(\hat{\lambda})^*$ test.

Size and power when a smooth transition occurs

Finally, Table 7 reports results for sizes of the tests when a *smooth transition* occurs under the null hypothesis, for the representative case of model 2a. The instant break tests are in general subject to very substantial size distortions. Moreover, further simulations involving larger sample sizes and the more general model 2 showed a general picture of severe oversizing. The only times the instant break tests reliably achieve close to correct size are when $\xi_i(\hat{\lambda})^*$ is applied to series which have medium to fast transitions of small magnitude.

As expected, the smooth transition tests are better sized. The conditional s_i^* test exhibits reasonable size behaviour for medium to fast transitions – being just a little undersized – although for slower transitions, this feature of undersizing can become quite severe. The unconditional s_i test is less reliable again, featuring moderate undersize or oversize in almost every case. These broad features persist with larger samples and also for model 2.

TABLE 7

Empirical sizes of nominal 5% level model 2a stationarity tests when a smooth transition occurs under the null: T = 100

δ_μ	γ'	λ'	$\tilde{\xi}_{2a}$	$\tilde{\xi}^*_{2a}$	$\xi_{2a}(\hat{\lambda})$	$\xi_{2a}(\hat{\lambda})^*$	s_{2a}	s^*_{2a}
2.5	1.5	0.25	8.90	0.11	5.39	1.70	3.51	0.54
		0.50	7.07	0.03	6.10	2.16	5.10	0.89
	3.0	0.25	36.73	4.16	16.15	8.80	2.15	0.61
		0.50	23.55	0.38	16.02	5.93	8.53	2.53
	15.0	0.25	37.80	5.21	13.99	8.24	4.19	1.37
		0.50	32.61	0.77	19.93	6.17	11.34	2.35
	30.0	0.25	28.15	2.92	9.51	5.36	6.30	2.24
		0.50	27.68	0.63	17.32	5.10	14.65	3.20
5.0	1.5	0.25	31.74	2.77	13.31	6.44	3.15	1.05
		0.50	14.09	0.19	10.63	4.12	7.93	1.76
	3.0	0.25	97.07	74.37	85.62	77.23	0.82	0.37
		0.50	71.90	5.14	46.84	20.09	6.39	3.44
	15.0	0.25	79.31	30.07	48.86	35.84	5.43	2.23
		0.50	66.01	2.10	38.05	12.87	12.34	2.28
	30.0	0.25	47.60	8.21	18.91	11.30	8.18	3.29
		0.50	38.62	0.71	22.53	6.83	16.61	3.52
10.0	1.5	0.25	95.61	65.10	80.31	68.94	1.47	0.74
		0.50	44.89	1.50	27.93	11.99	15.94	7.30
	3.0	0.25	100.00	100.00	100.00	100.00	0.20	0.10
		0.50	99.97	79.49	98.83	86.26	1.00	0.23
	15.0	0.25	99.93	94.88	98.32	95.69	4.93	2.44
		0.50	99.60	27.06	89.87	46.16	11.94	2.18
	30.0	0.25	86.51	34.54	54.92	40.54	8.51	3.65
		0.50	75.77	1.51	41.44	11.22	16.69	3.54

Notes: See notes to Tables 1, 2 and 4 for model and test definitions, and Table 3 for parameter definitions. δ_μ denotes the magnitude of the smooth transition in level.

When the smooth transition tests are applied to stationary series with *instantaneous* breaks, unreported simulations indicate that $\hat{\gamma}$ diverges (as expected) and $\hat{\lambda}$ superconsistently estimates the break fraction (evidence of convergence at rate *T*), leading to the correctly sized conditional s_i^* test observed in Table 4. However, when the data-generating process (DGP) contains a *smooth transition*, as in Table 7, the transition parameter estimates are still consistent, but simulation indicates convergence only at rate $\sqrt{T}$. This slower convergence explains the fact that size distortions for the conditional test s_i^* remain even in very large samples when a transition is present: for speeds of $\gamma' = 15$ and $\gamma' = 30$, the test remains a little undersized, while for slower speeds, more severe distortions are observed, predominantly those of undersize. The undersize results from errors in the estimation of λ, often leading to the removal of any appearance of a unit root. Compared with the other tests, however, it is apparent that s_i^* is the only test that has, irrespective of the sample size and magnitude of structural change, reasonable properties for models with moderate to fast transitions. However, none of the tests can be considered satisfactory for models involving slower transitions.

Regarding power comparisons when a smooth transition occurs under the alternative hypothesis, the instant break tests will always have a tendency to outperform the smooth transition tests. This follows in part from the former being oversized when smooth transitions occur under the null, but also from the way the stationarity tests are conducted. The tests are based on testing for stationarity in the residuals from a fitted deterministic model; thus, if the true process is unit root with a smooth transition, the residuals from an instant break model will generally appear less stationary than those from a smooth transition model, and so the instant break tests will tend to reject more often. The main concern with stationarity testing when smooth transitions in linear trend are present is one of size behaviour, as analysed above. However, unreported simulations confirm that the smooth transition tests s_i and s_i^* are consistent under the alternative hypothesis.

V. Empirical applications

To illustrate the application of these tests, we investigate the trend behaviour of four economic time series. These are the end-year value of the *S&P 500* stock market index for the period 1870–2000, annual UK gross domestic product from 1855–1999, the annual British index of industrial production from 1700–1913, and end-month observations on the dollar–sterling exchange rate from January 1988 to December 2001. The logarithms of the first three series are analysed, but no transformation is made on the exchange rate.

The stationarity tests in this paper are applied to the series, with the exception of $\tilde{\xi}_i^*$ because of the test's poor size and power performance when structural change is present. In order to account for potential residual autocorrelation when conducting the tests, the KPSS nonparametric modification was employed, i.e. $\hat{\sigma}^2$ in the test statistics was replaced with $s^2(l)$ where

$$s^2(l) = T^{-1}\sum_{t=1}^{T} e_t^2 + 2T^{-1}\sum_{j=1}^{l} w(j,l)\sum_{t=j+1}^{T} e_t e_{t-j}. \tag{15}$$

Following KPSS, we use the Bartlett window for the weighting function: $w(j, l) = 1 - j/(l + 1)$, with the lag truncation parameter chosen by $l = [4(T/100)^{1/4}]$ where $[\cdot]$ denotes the integer part (see, e.g. Schwert, 1989). Critical values can be taken from Tables 1–3 for the $\xi_i(\hat{\lambda})$, s_i and s_i^* tests, and from Harvey and Mills (2002) and Busetti and Harvey (2001) for the $\tilde{\xi}_i$ and

TABLE 8

Empirical applications of stationarity tests

	$\tilde{\xi}_i$	$\xi_i(\hat{\lambda})$	$\xi_i(\hat{\lambda})^*$	s_i	s_i^*
(A) S&P 500 index					
Model 1	0.534***	0.607***	0.607***	0.069	0.069
	$(\hat{\tau} = 1948)$	$(\hat{\tau} = 1954)$		$(\hat{\lambda} = 6.200, \hat{\gamma} = 0.771)$	
Model 2a	0.228***	0.265***	0.265***	0.068	0.068
	$(\hat{\tau} = 1966)$	$(\hat{\tau} = 1912)$		$(\hat{\lambda} = 7.545, \hat{\gamma} = 0.748)$	
Model 2b	0.057	0.059	0.059		
	$(\hat{\tau} = 1945)$	$(\hat{\tau} = 1943)$			
Model 2	0.056***	0.059*	0.059*	0.063*	0.063*
	$(\hat{\tau} = 1935)$	$(\hat{\tau} = 1939)$		$(\hat{\lambda} = 0.617, \hat{\gamma} = 44.61)$	
(B) UK GDP					
Model 1	0.779***	0.852***	0.852***	0.217*	0.217***
	$(\hat{\tau} = 1928)$	$(\hat{\tau} = 1939)$		$(\hat{\lambda} = 28.48, \hat{\gamma} = 0.147)$	
Model 2a	0.137***	0.282***	0.282***	0.293***	0.293***
	$(\hat{\tau} = 1912)$	$(\hat{\tau} = 1918)$		$(\hat{\lambda} = 0.448, \hat{\gamma} = 76.25)$	
Model 2b	0.147***	0.157***	0.157***		
	$(\hat{\tau} = 1958)$	$(\hat{\tau} = 1950)$			
Model 2	0.036	0.036	0.036	0.031	0.031
	$(\hat{\tau} = 1919)$	$(\hat{\tau} = 1919)$		$(\hat{\lambda} = 0.451, \hat{\gamma} = 83.06)$	
(C) British industrial production					
Model 1	0.990***	1.155***	1.155***	0.160	0.160*
	$(\hat{\tau} = 1820)$	$(\hat{\tau} = 1832)$		$(\hat{\lambda} = 0.679, \hat{\gamma} = 1.934)$	
Model 2a	0.434***	0.473***	0.473***	0.089**	0.089**
	$(\hat{\tau} = 1752)$	$(\hat{\tau} = 1744)$		$(\hat{\lambda} = 0.678, \hat{\gamma} = 2.563)$	
Model 2b	0.195***	0.210***	0.210***		
	$(\hat{\tau} = 1787)$	$(\hat{\tau} = 1791)$			
Model 2	0.190***	0.215***	0.215***	0.039	0.039
	$(\hat{\tau} = 1780)$	$(\hat{\tau} = 1791)$		$(\hat{\lambda} = 0.591, \hat{\gamma} = 3.841)$	
(D) Dollar–sterling exchange rate					
Model 1	0.137**	0.179	0.179*	0.184	0.184*
	$(\hat{\tau} = 1993{:}5)$	$(\hat{\tau} = 1992{:}9)$		$(\hat{\lambda} = 0.345, \hat{\gamma} = 1195.1)$	
Model 2a	0.088***	0.182***	0.182***	0.100**	0.100*
	$(\hat{\tau} = 1996{:}5)$	$(\hat{\tau} = 1992{:}10)$		$(\hat{\lambda} = 0.625, \hat{\gamma} = 15.02)$	
Model 2b	0.122***	0.123***	0.123***		
	$(\hat{\tau} = 1993{:}8)$	$(\hat{\tau} = 1994{:}2)$			
Model 2	0.081***	0.167***	0.167***	0.082**	0.082**
	$(\hat{\tau} = 1996{:}11)$	$(\hat{\tau} = 1992{:}9)$		$(\hat{\lambda} = 0.680, \hat{\gamma} = 6.926)$	

Notes: See notes to Tables 1, 2 and 4 for model and test definitions. *, ** and *** denote rejection at the 10%, 5% and 1% levels, respectively.

$\xi_i(\hat{\lambda})^*$ tests, respectively. In order to achieve a little more accuracy, we actually generated critical values by simulation for the exact sample sizes of the four series, using the KPSS correction in the tests, and for finer grids in the case of the conditional procedures.

The *S&P 500* series is an update of that used by Perron (1989) in his seminal study of testing for stationarity in the presence of trend breaks, in which he was able to reject the null of a unit root in favour of stationary deviations about a trend function with breaks in both level and slope at 1929. The results of applying the various tests to the four possible trend break models are shown in panel A of Table 8. For models 1 and 2a, there is a clear rejection of stationarity (model 1) or trend stationarity (model 2a) around an instantaneous break in level, but no rejection when a smooth transition in level is assumed. A glance at the first and second rows of Figure 1, which superimposes each of the fitted trends in turn onto the observed series, easily explains these results. The simple instant-level break models are clearly inappropriate trend specifications, so that the deviations from such trends are obviously nonstationary. The smooth transition models, however, provide much better fits to the series, and do not allow a rejection of the stationary null. However, the parameter estimates imply extremely slow transitions with a midpoint estimated to be well outside the sample period. In these circumstances, the tests are substantially undersized, and what is being fit is basically equivalent to a low order polynomial in time. The piecewise trend model 2b also does not allow rejection of the null, with the tests being almost identical in selecting the break point to be in the mid-1940s (see the third row of Figure 1). Allowing both trend and level shifts also provides little evidence against the null, with only a rejection at the 10% level using the preferred $\xi_2(\hat{\lambda}), \xi_2(\hat{\lambda})^*$ tests allowing for an instant trend and level break, and a similar result when a smooth transition in level and trend is assumed. The strong rejection implied by $\tilde{\xi}_2$ is most likely due to the substantial oversizing that the test displays for breaks of this magnitude, as observed in the simulations. We would thus conclude that, if we are to model the *S&P 500* index as stationary deviations about a deterministic trend, then a formulation allowing an evolving trend is necessary, as found by Perron (1989) for a sample period ending in 1970. Our examination of a longer sample places the break date rather later than the 1929 assumed by Perron.

The tests for UK output are reported in panel B, and show clear rejections for all models except model 2, indicating that the series can be modelled as stationary about a change in level and trend. Some interesting features emerge from these models. Allowing shifts in both level and trend fits the break at 1919, with the corresponding smooth transition having a rapid transition with a midpoint also during 1919. The underlying models estimate the pre- and

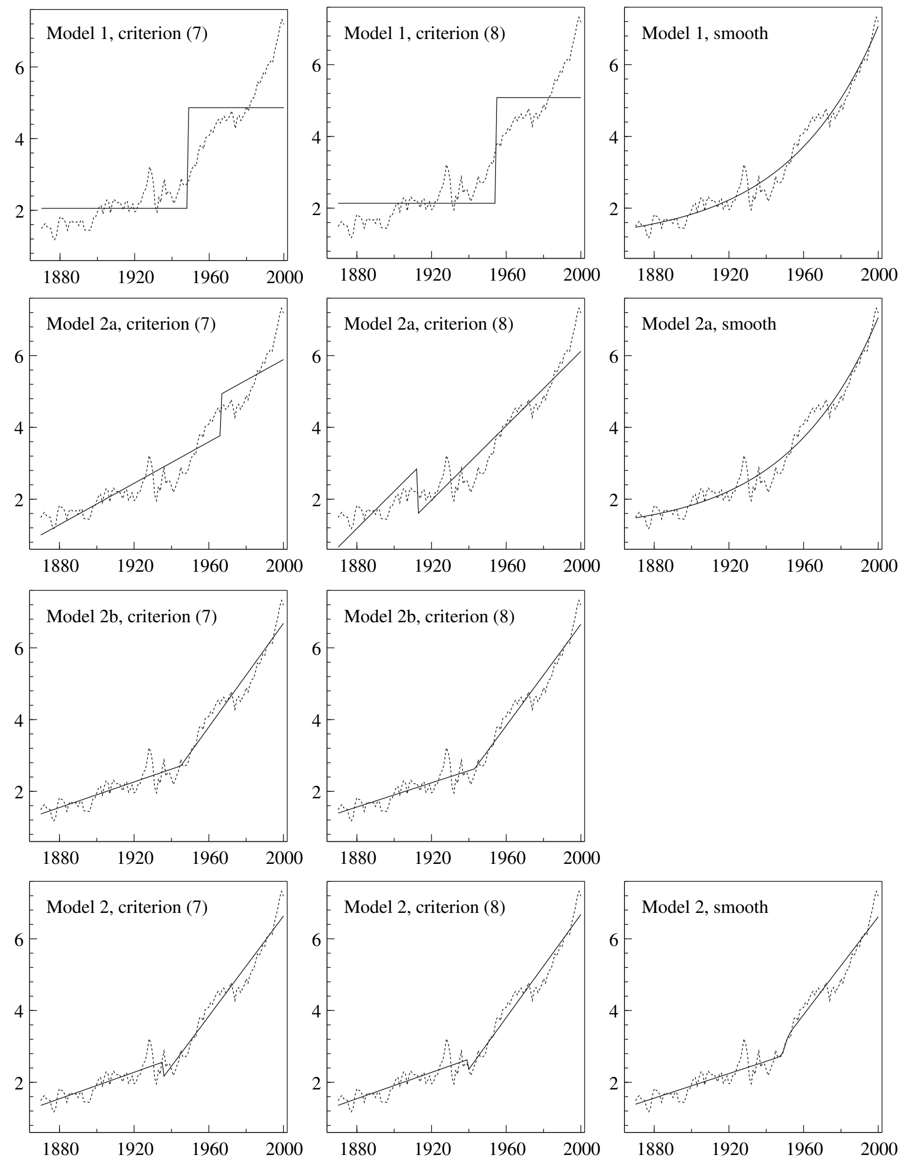

Figure 1. *S&P 500* index and fitted trend functions

post-break trend growth rates to be 1.9% and 2.2% per annum, respectively, and these are consistent with a precursor model presented in Mills (1994). Model 2a restricts the pre- and post-break trend growths to be identical, allowing only level shifts in output. It is clear from the second row of Figure 2 that break date estimation using criterion (8) provides a much more plausible trend function than that suggested by criterion (7), illustrating the general superiority of equation (8) over (7). The imposition of the pre- and post-break trend growth restriction, which is difficult to detect visually in Figure 2 but is statistically rejected in model 2, is responsible for the rejection of stationarity in model 2a. This underlines the sensitivity of the tests to the correct trend specification.

Panel C reports the results for the index of industrial production, whose trend evolution has been analysed in the context of the industrial revolution by Mills and Crafts (1996) and Crafts and Mills (1997). In the latter, a model 2-type smooth transition was fitted after a rejection of the unit root null hypothesis had been found using the Leybourne *et al.* (1998) testing procedure. The s_2 and s_2^* tests are in agreement with this view that deviations in industrial production from a smooth transition in linear trend are stationary (although the reliability of the tests for the estimated transition speed is somewhat questionable), while all the instantaneous break tests reject stationarity; these models have obviously inferior fits, as can be seen in Figure 3.

Finally, panel D reports the test statistics for the dollar–sterling exchange rate. The sample period was chosen to start after a period of adjustment following the Louvre Accord in March 1987. Here the test statistics follow a rather different pattern to those of the previous examples. Clear rejections are found for all statistics except the preferred instant break and smooth transition versions for model 1. The reasons for this are again made clear in Figure 4. All models containing trends clearly lead to unrealistic trend functions and nonstationary residuals. In addition, model 1 where the break date is estimated using equation (7) also results in an unrealistic trend function, placing the break somewhat later than that implied by equation (8) and the smooth transition model, again confirming the relative advantage of equation (8) over (7). The acceptable statistics are thus for those models that embody just an instantaneous level shift, or equivalently, an extremely fast smooth transition in mean. The break point is identified to be at September 1992, when the UK abruptly left the European Exchange Rate Mechanism (ERM). It would thus seem that the dollar–sterling exchange rate might be characterized as stationary deviations about a constant mean that underwent an abrupt downward shift in response to a regime change. However, investigation of the stochastic properties of the deviations in the two regimes offers further insight into the behaviour of this exchange rate. During the ERM regime, the deviations are characterized by a second order autoregressive process with complex roots of $0.62 \pm 0.10i$, thus producing the 'long swings' around a constant mean of 1.76 observed in the data. Although the ERM did not set a band for the dollar–sterling exchange rate, it seems to have constrained the fluctuations that this rate could take. The

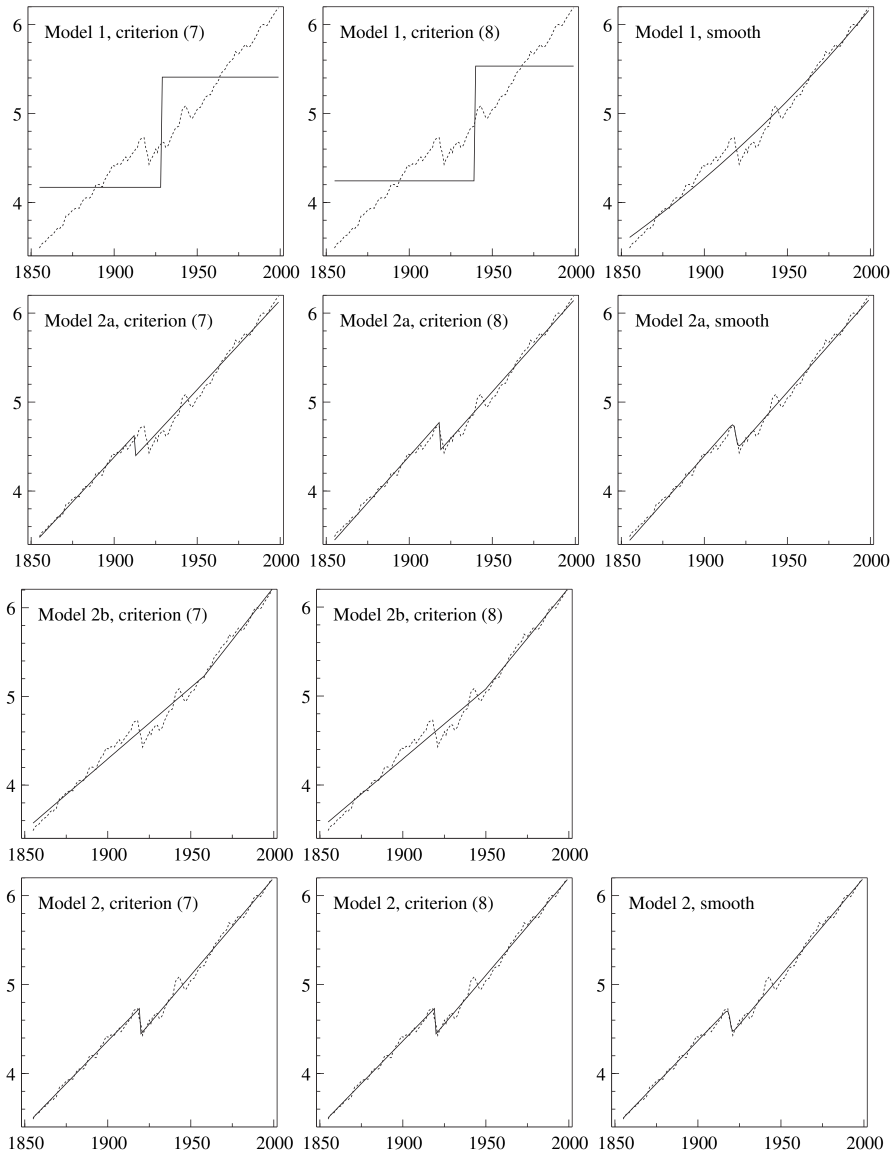

Figure 2. UK GDP and fitted trend functions

process for the post-ERM period is best represented by a first order autoregression with a root of 0.92: in other words, it is close to the driftless random walk that is usually found to characterize freely floating exchange rate regimes, perhaps giving rise to the rejections at the 10% level observed for $\xi_1(\hat{\lambda})^*$ and s_1^*.

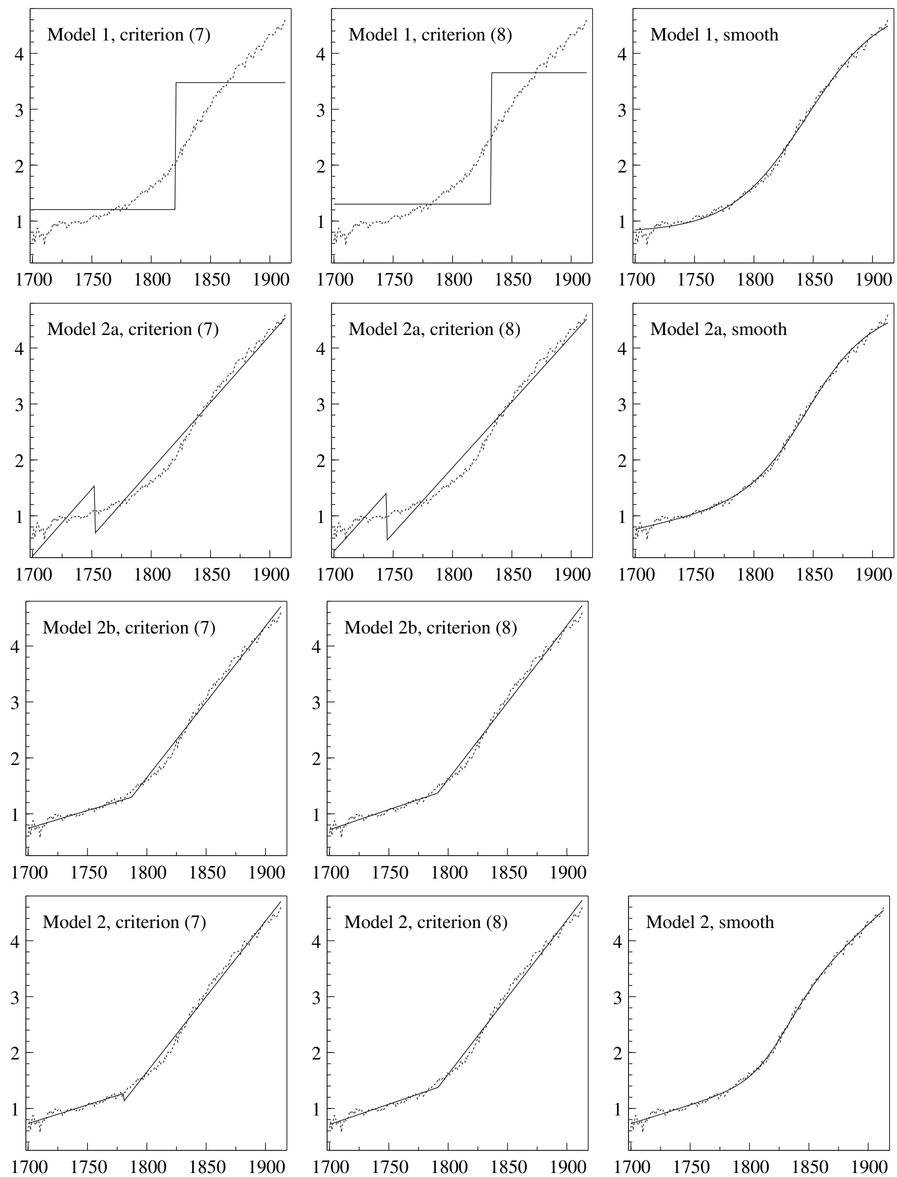

Figure 3. British industrial production and fitted trend functions

VI. Conclusion

We have examined several tests of the null hypothesis of stationarity against a unit root alternative where the series concerned admits structural change at an unknown point in time. Four tests are considered which permit an

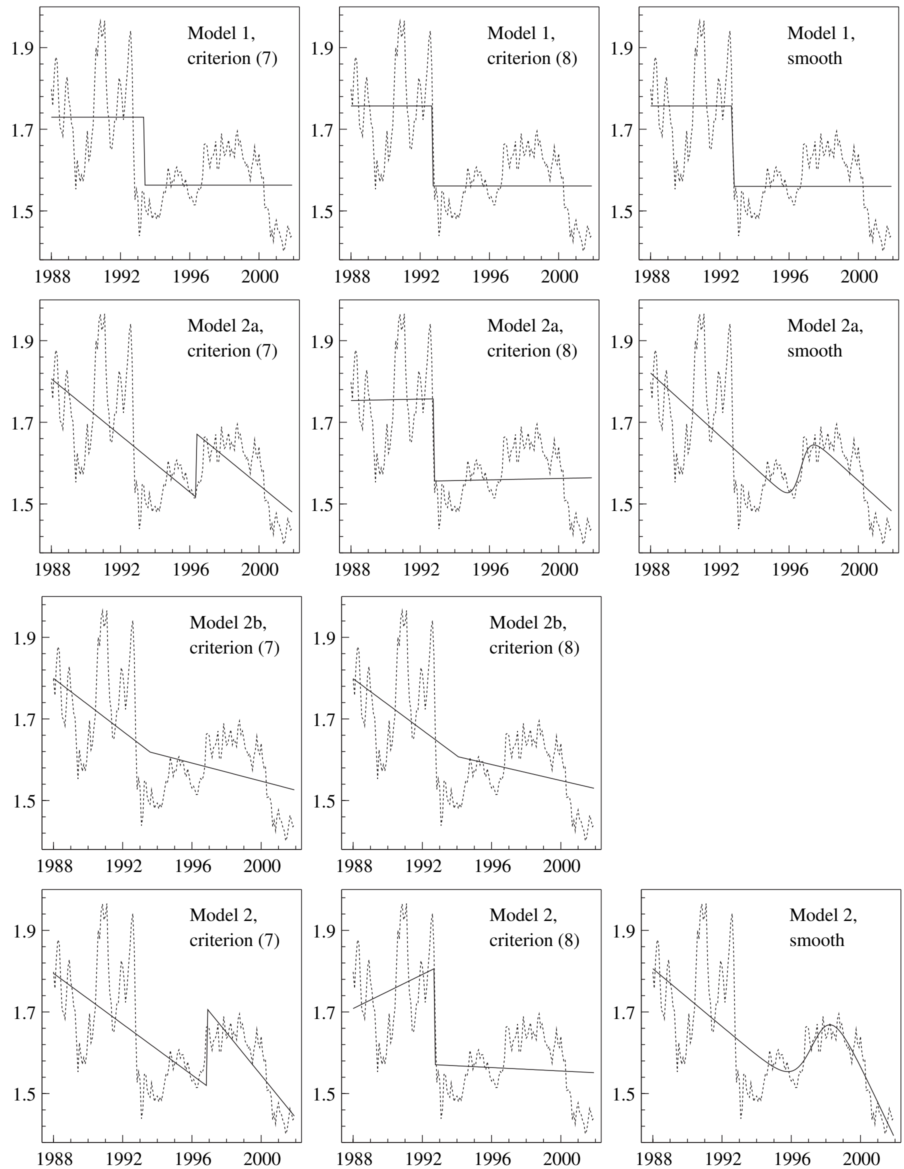

Figure 4. Dollar–sterling exchange rate and fitted trend functions

instantaneous break in level and/or trend, and two which allow a more general deterministic structure where the series undergoes a smooth transition between two different level and/or trend regimes.

The instant break tests rely, at least implicitly, on some method for estimating the unknown break date. The simulation results show clearly that break date estimation based on the significance of relevant dummy variable coefficients (or, equivalently, minimum sum of squared residuals or SBC) consistently outperforms Zivot-Andrews-type approaches where the break date is selected to give most weight to the null of (trend) stationarity about a break. Tests based on the latter approach are inferior to the former in terms of size distortions when a break occurs under the null. Of the two procedures that employ the preferred break date estimator, the conditional test, using critical values derived from tests with a known break fraction, outperforms the unconditional test based on the shift assumption. The results highlight the inadequacy of the shift assumption for breaks of interesting magnitude, and the reliable size of the conditional test. We therefore recommend using the conditional approach where dummy variable coefficient significance is used to determine the break date; this procedure coincides with that proposed by Kurozumi (2002), Busetti and Harvey (2002) and Lee (1999). The preferred test has less power than alternative procedures in some circumstances, but generally the cost in terms of foregone power from using such a reliably sized test is small.

A possible caveat to this recommendation exists when no break is actually present in the series, since the preferred test suffers from substantial undersizing and low power in these circumstances. Busetti and Harvey (2002) argue that if uncertainty exists concerning the presence of a break, their unconditional Zivot-Andrews-type test should instead be used. On the basis of our results, we would modify this recommendation to suggest use of the newly proposed variant of this test, maintaining the shift assumption for determining the critical values, but using the dummy variable coefficient break date estimator. This follows because, if a small break does occur in the series, the size distortions are less severe when using the superior estimator of the break timing. Our recommendation is therefore always to use tests where the break date is selected on the basis of dummy variable coefficient significance, but modifying the critical values depending on the degree of confidence in the existence of a break. This approach is also appealing in the sense that break date selection is then always consistent with the preferred method in the parallel context of testing the unit root null against a stationary alternative.

The new smooth transition tests provide additional generality to the stationarity testing procedure. The preferred version of the test, which is conditional on the estimated transition parameters and uses critical values associated with tests where the transition midpoint and speed are known, has good size properties and reasonable power when an instant break occurs under the null, and has greatly superior size performance compared with the instant break tests if the null is of stationarity around a moderate to fast smooth transition. The limitations of the procedure are when the transition speed is slow, or when no break or transition occurs. In the latter case, the unconditional smooth transition test has correct size by construction, thus an argument (equivalent to that above for instantaneous break procedures) could be made for using the unconditional approach if uncertainty exists regarding the presence of a transition.

Given that structural change in economic time series often occurs gradually rather than instantaneously, these smooth transition tests have potential value in practical applications and provide a useful stationarity testing counterpart to the unit root tests proposed by Leybourne *et al.* (1998). The choice of when to employ an instantaneous break or smooth transition test must be guided by the application. If instantaneous structural change is likely, both sets of tests are applicable, while if a transition of moderate to fast speed is more apparent, only the smooth transition tests should be considered. If a slow smooth transition is detected, the question is raised as to the reliability of the tests, and, as highlighted in the empirical applications, this may suggest that a trend or polynomial trend specification may be more appropriate.

Final Manuscript Received: May 2004

References

Bacon, D. W. and Watts, D. G. (1971). 'Estimating the transition between two intersecting straight lines', *Biometrika*, Vol. 58, pp. 525–534.

Busetti, F. and Harvey, A. C. (2001). 'Testing for the presence of a random walk in series with structural breaks', *Journal of Time Series Analysis*, Vol. 22, pp. 127–150.

Busetti, F. and Harvey, A. C. (2002). 'Further comments on stationarity tests in series with structural breaks at unknown points', *Journal of Time Series Analysis*, Vol. 24, pp. 137–140.

Crafts, N. F. R. and Mills, T. C. (1997). 'Endogenous innovation, trend growth and the British industrial revolution', *Journal of Economic History*, Vol. 57, pp. 950–956.

Dickey, D. A. and Fuller, W. A. (1979). 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association*, Vol. 73, pp. 427–431.

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

Harvey, D. I., Leybourne, S. J. and Newbold, P. (2001). 'Innovational outlier unit root tests with an endogenously determined break in level', *Oxford Bulletin of Economics and Statistics*, Vol. 63, pp. 559–575.

Harvey, D. I. and Mills, T. C. (2002). 'A note on Busetti–Harvey tests for stationarity in series with structural breaks', *Journal of Time Series Analysis*, Vol. 24, pp. 159–164.

Kurozumi, E. (2002). 'Testing for stationarity with a break', *Journal of Econometrics*, Vol. 108, pp. 63–99.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992). 'Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?', *Journal of Econometrics*, Vol. 54, pp. 159–178.

Lee, J. (1999). 'Stationarity tests with multiple endogenized breaks', in Rothman P. (ed.), *Nonlinear Time Series Analysis of Economic and Financial Data*, Kluwer Academic Press, Massachusetts.

Lee, J. and Strazicich, M. (2001). 'Testing the null of stationarity in the presence of a structural break', *Applied Economics Letters*, Vol. 8, pp. 377–382.

Leybourne, S. J. and McCabe, B. P. M. (1994). 'A consistent test for a unit root', *Journal of Business and Economic Statistics*, Vol. 12, pp. 157–166.

Leybourne, S. J. and McCabe, B. P. M. (1999). 'Modified stationarity tests with data-dependent model-selection rules', *Journal of Business and Economic Statistics*, Vol. 17, pp. 264–270.

Leybourne, S. J., Newbold, P. and Vougas, D. (1998). 'Unit roots and smooth transitions', *Journal of Time Series Analysis*, Vol. 19, pp. 83–97.

Lin, C. J. and Teräsvirta, T. (1994). 'Testing the constancy of regression parameters against continuous structural change', *Journal of Econometrics*, Vol. 62, pp. 211–228.

MacNeill, I. B. (1978). 'Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times', *Annals of Statistics*, Vol. 6, pp. 422–433.

Mills, T. C. (1994). 'Segmented trends and the stochastic properties of UK output', *Applied Economics Letters*, Vol. 1, pp. 132–133.

Mills, T. C. and Crafts, N. F. R. (1996). 'Trend growth in British industrial output, 1700–1913: a reappraisal', *Explorations in Economic History*, Vol. 33, pp. 277–295.

Nyblom, J. and Mäkeläinen, T. (1983). 'Comparisons of tests for the presence of random walk coefficients in a simple linear model', *Journal of the American Statistical Association*, Vol. 78, pp. 856–864.

Perron, P. (1989). 'The Great Crash, the oil price shock, and the unit root hypothesis', *Econometrica*, Vol. 57, pp. 1361–1401.

Perron, P. (1993). 'The Great Crash, the oil price shock, and the unit root hypothesis: Erratum', *Econometrica*, Vol. 61, pp. 248–249.

Schwert, G. W. (1989). 'Tests for unit roots: a Monte Carlo investigation', *Journal of Business and Economic Statistics*, Vol. 7, pp. 147–159.

Sollis, R., Leybourne, S. J. and Newbold, P. (1999). 'Unit roots and asymmetric smooth transitions', *Journal of Time Series Analysis*, Vol. 20, pp. 671–677.

Vogelsang, T. J. and Perron, P. (1998). 'Additional tests for a unit root allowing for a break in the trend function at an unknown time', *International Economic Review*, Vol. 39, pp. 1073–1100.

Zivot, E. and Andrews, D. W. K. (1992). 'Further evidence on the Great Crash, the oil price shock, and the unit-root hypothesis', *Journal of Business and Economic Statistics*, Vol. 10, pp. 251–270.

Appendix

Proof of equation (9): In order to prove that the asymptotic distribution of $\xi_i(\hat{\lambda})$ $(i = 1, 2, 2\text{a}, 2\text{b})$ is as given in equation (9), it is sufficient to show that the limiting distributions of $t_{\hat{\delta}_\mu}$ for models 1 and 2a, $t_{\hat{\delta}_\beta}$ for model 2b, and $F_{\hat{\delta},\hat{\beta}}$ for model 2, are given by $Z_i(\lambda)$ ($i = 1$, 2a, 2b, 2) respectively. The result then follows from application of the continuous mapping theorem and the established asymptotic results of Busetti and Harvey (2001) and Harvey and Mills (2002) for the stationarity test statistic with a given break fraction.

Consider first model 1 (proofs for models 2a and 2b follow with appropriate modifications and are omitted). Under the null of stationarity $(\sigma_\eta^2 = 0)$ about a break in level of magnitude δ_μ at time τ', the generating process can be written as follows, assuming $\mu_0 = 0$ without loss of generality:

$$y_t = \delta_\mu w_t' + \varepsilon_t.$$

The appropriate regression for this model with an assumed break date τ is then

$$y_t = \mu + \delta_\mu w_t + \varepsilon_t$$

with the ordinary least squares (OLS) estimator of δ_μ given by

$$\hat{\delta}_\mu = (T-\tau)^{-1}\sum_{t=\tau+1}^{T} y_t - \tau^{-1}\sum_{t=1}^{\tau} y_t.$$

Substituting in for y_t and scaling appropriately yields

$$T^{1/2}\hat{\delta}_\mu = T^{1/2}\delta_\mu\left[\frac{1-\lambda'}{1-\lambda} - I\frac{\lambda-\lambda'}{\lambda(1-\lambda)}\right] + T^{-1/2}(1-\lambda)^{-1}\sum_{t=\tau+1}^{T}\varepsilon_t - T^{-1/2}\lambda^{-1}\sum_{t=1}^{\tau}\varepsilon_t$$

where $I = 1(\tau > \tau')$ and $1(\cdot)$ is the indicator function. Now the shift assumption can be represented as $\delta_\mu = kT^{-(\kappa+\frac{1}{2})}$ where $\kappa > 0$. Use of this representation, combined with Lemma 2 of Busetti and Harvey (2001), leads to the asymptotic distribution of $\hat{\delta}_\mu$:

$$\sigma^{-1}T^{1/2}\hat{\delta}_\mu \Rightarrow (1-\lambda)^{-1}[W(1)-W(\lambda)] - \lambda^{-1}W(\lambda).$$

The estimated variance of $\hat{\delta}_\mu$ is given by

$$\hat{\sigma}^2_{\hat{\delta}_\mu} = \hat{\sigma}^2\tau^{-1}(T-\tau)^{-1}T.$$

Scaling, and noting that $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$, gives the limiting behaviour of this estimator:

$$T\hat{\sigma}^2_{\hat{\delta}_\mu} \xrightarrow{p} \sigma^2\lambda^{-1}(1-\lambda)^{-1}.$$

The asymptotic distribution of $t_{\hat{\delta}_\mu} = \hat{\delta}_\mu/\hat{\sigma}_{\hat{\delta}_\mu}$ then follows:

$$t_{\hat{\delta}_\mu} \Rightarrow \sqrt{\frac{\lambda}{1-\lambda}}\left[W(1) - \frac{1}{\lambda}W(\lambda)\right].$$

Consider now model 2. The generating process under the null can be expressed as follows, assuming $\mu_0 = 0$ and $\beta = 0$ without loss of generality:

$$y_t = \delta_\mu w_t' + \delta_\beta(w_t' t) + \varepsilon_t.$$

The appropriate regression with an assumed break date τ is now

$$y_t = \mu + \beta t + \delta_\mu w_t + \delta_\beta (w_t t) + \varepsilon_t$$

with the (scaled) OLS estimators of the regression parameters given by

$$\begin{bmatrix} T^{1/2}\hat{\mu} \\ T^{3/2}\hat{\beta} \\ T^{1/2}\hat{\delta}_\mu \\ T^{3/2}\hat{\delta}_\beta \end{bmatrix} \approx \begin{bmatrix} 1 & \frac{1}{2} & 1-\lambda & \frac{1}{2}(1-\lambda^2) \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{2}(1-\lambda^2) & \frac{1}{3}(1-\lambda^3) \\ 1-\lambda & \frac{1}{2}(1-\lambda^2) & 1-\lambda & \frac{1}{2}(1-\lambda^2) \\ \frac{1}{2}(1-\lambda^2) & \frac{1}{3}(1-\lambda^3) & \frac{1}{2}(1-\lambda^2) & \frac{1}{3}(1-\lambda^3) \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2}\sum_{t=1}^{T} y_t \\ T^{-3/2}\sum_{t=1}^{T} t y_t \\ T^{-1/2}\sum_{t=\tau+1}^{T} y_t \\ T^{-3/2}\sum_{t=\tau+1}^{T} t y_t \end{bmatrix}$$

in large samples. Lemma 2 of Harvey and Busetti (2001) and the shift assumption, which here implies $\delta_\mu = k_1 T^{-(\kappa_1+\frac{1}{2})}$ and $\delta_\beta = k_2 T^{-(\kappa_2+\frac{3}{2})}$, where κ_1, $\kappa_2 > 0$, allows derivation of the limits of the final right hand side terms:

$$T^{-1/2}\sum_{t=1}^{T} y_t \Rightarrow \sigma W(1)$$

$$T^{-3/2}\sum_{t=1}^{T} t y_t \Rightarrow \sigma \int_0^1 r \,\mathrm{d}W(r)$$

$$T^{-1/2}\sum_{t=\tau+1}^{T} y_t \Rightarrow \sigma[W(1) - W(\lambda)]$$

$$T^{-3/2}\sum_{t=\tau+1}^{T} t y_t \Rightarrow \sigma \int_\lambda^1 r \,\mathrm{d}W(r).$$

The asymptotic distributions of the relevant parameter estimates then follow:

$$T^{1/2}\hat{\delta}_\mu \Rightarrow 2\sigma\lambda^{-2}(1-\lambda)^{-3}H_1$$

$$T^{3/2}\hat{\delta}_\beta \;\Rightarrow\; 6\sigma\lambda^{-3}(1-\lambda)^{-3}H_2$$

where H_1 and H_2 are as defined below equation (9). Now

$$F_{\hat{\delta},\hat{\beta}} = \frac{(Rb)'[R(X'X)^{-1}R']^{-1}(Rb)}{2\hat{\sigma}^2}$$

where b is the vector of OLS parameter estimates, X is the matrix of regressors, and

$$R = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This allows us to write

$$F_{\hat{\delta},\hat{\beta}} \approx \begin{bmatrix} T^{1/2}\hat{\delta}_\mu & T^{3/2}\hat{\delta}_\beta \end{bmatrix} \begin{bmatrix} \frac{4(1-2\lambda+4\lambda^2)}{\lambda(1-\lambda)^3} & \frac{-6(1-3\lambda+4\lambda^2)}{\lambda^2(1-\lambda)^3} \\ \frac{-6(1-3\lambda+4\lambda^2)}{\lambda^2(1-\lambda)^3} & \frac{12(1-3\lambda+3\lambda^2)}{\lambda^3(1-\lambda)^3} \end{bmatrix}^{-1} \begin{bmatrix} T^{1/2}\hat{\delta}_\mu \\ T^{3/2}\hat{\delta}_\beta \end{bmatrix} / 2\hat{\sigma}^2.$$

Given $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$, the asymptotic distribution of $F_{\hat{\delta},\hat{\beta}}$ can then be obtained:

$$F_{\hat{\delta},\hat{\beta}} \Rightarrow \frac{2}{\lambda^3(1-\lambda)^5}[(1-3\lambda+3\lambda^2)H_1^2+3(1-3\lambda+4\lambda^2)^2 H_1 H_2+3(1-2\lambda+4\lambda^2)H_2^2].$$