Innovational outlier unit root tests with an endogenously determined break in level

DAVID I. HARVEY, STEPHEN J. LEYBOURNE and PAUL NEWBOLD

Department of Economics, Loughborough University
School of Economics, University of Nottingham
School of Economics, University of Nottingham

I. Introduction

Structural breaks in trend can strongly influence Dickey-Fuller-type unit root tests, and tests that are valid in the presence of such a break at a known point in time have been developed by Perron (1989, 1990, 1993) and Perron and Vogelsang (1992b). Tests are available to allow for a break in level or slope, or both, for breaks of ‘additive outlier’ (AO) type, where the break is taken to be sudden, and the ‘innovational outlier’ (IO) type, where the break is allowed to evolve over time. However, as argued for example by Christiano (1992), such tests are not appropriate in circumstances where the analyst is prompted to select the break date by reference to the data, such as through graphical inspection. It is therefore also desirable to have available tests in which that date is not assumed known a priori, but is rather treated as endogenous. Such tests have been proposed and analysed by, among others, Banerjee et al. (1992), Perron and Vogelsang (1992a), Zivot and Andrews (1992), Perron (1997), and Vogelsang and Perron (1998). These tests are generally based on one of two approaches, both of which involve fitting the appropriate Dickey-Fuller-style regressions, augmented with dummy variables, for all possible break dates. One approach is to choose the break date for which the Dickey-Fuller $t$-statistic is a minimum, that is least favourable to the unit root null hypothesis. The other is to choose the date that seems ‘most likely’ on the basis of test statistics associated with coefficient estimates on the relevant break dummy variables.

It would seem preferable that endogenous break tests be valid when there is a break under the null hypothesis, a point stressed by Nunes et al. (1997) and extensively analysed by Vogelsang and Perron (1998). When there is a
break in slope or drift, whether or not there is also a break in level, the theoretical results of these latter authors are mixed, and not entirely reassuring. It is shown that, for non-zero breaks in drift, the sizes of tests based on least favourable Dickey-Fuller-type \( t \)-statistics approach one as the number of sample observations grows, so that for very large samples spurious rejections of the null hypothesis become almost certain. This suggests that one should concentrate instead on tests based on the apparent strength of evidence for particular break dates implied by estimated coefficients on the dummy variables. In the AO case, Vogelsang and Perron (1998) show that such approaches select the correct break date asymptotically, and hence that the associated unit root test statistics have the same limiting null distribution as in the case where the break date is known. Unfortunately, this differs from the limiting null distribution when there is no break. However, as a practical matter the analyst may be unwilling to assume knowledge of whether there is a break under the null, or to be convinced by an argument that seems to suggest that in finite sample applications one set of critical values should be used for the no break case, and an entirely different set when there is a break, however miniscule its magnitude. Recognising this dilemma, Vogelsang and Perron (1998) note that a conservative approach is to employ critical values associated with the no break case, leading to tests that are somewhat undersized when there is in fact a break. In the case of a break in slope, Vogelsang and Perron show that break date selection based on test statistics associated with estimated coefficients on dummy variables does not for IO models lead asymptotically to correct choice so that size distortions of associated unit root tests can occur asymptotically when there is in fact a break under the null hypothesis.

On the surface, as argued by Perron and Vogelsang (1992a) and Vogelsang and Perron (1998), the position is considerably more straightforward when there is a break only in level, irrespective of whether or not a linear trend term with unbroken slope is included in the model. Specifically, the test statistics are all asymptotically invariant to the magnitude of the break and critical values can be developed from the model with no break under the null hypothesis.

We argue in the remainder of this paper that such reassurance can prove more illusory than real. Simulation evidence is presented to show that in reasonable-sized samples IO, but not AO, tests where break date selection is based on the significance of coefficients on dummy variables can lead to strong spurious rejections of the unit root hypothesis when there is a large break under the null. The same is true for both the IO and AO case when selection is based on the least favourable Dickey-Fuller \( t \)-statistic. The first of these cases is explored in some detail, since dummy variable significance appears in general to be the less unsatisfactory of the two approaches. We
note that the asymptotic analysis that leads to a limiting null distribution independent of break magnitude is based on holding that magnitude constant as sample size increases. However, as argued by Leybourne and Newbold (2000) in the context of Dickey-Fuller-type tests and Kim et al. (2000) in the context of Perron tests, an asymptotic analysis in which break magnitude is proportional to the square root of sample size can often generate more reliable predictions of what will occur for series of practically interesting lengths. Using such a normalisation, we show that in the IO case a break date different from the true date is likely to be chosen through the $t$-ratios on estimated coefficients associated with the level break dummy variable. Moreover, this chosen break date, immediately before the correct date, is in precisely the place shown by Kim et al. (2000) to lead to most severe spurious rejections of the unit root null hypothesis. This analysis suggests a small modification of the test which, as will be seen, avoids the spurious rejection problem while retaining satisfactory power under the alternative hypothesis.

Lee and Strazicich (1999) also examined, entirely through simulation, the IO case, finding the estimated break date and spurious rejection problem for large breaks. These authors also included the possibility of a simultaneous break in level and slope, though, as we argue, it is the former that is the source of difficulty.

II. Tests Allowing a Break in Level

Let the time series $y_t$ be generated by the model

$$y_t = \theta D(T_B)_t + y_{t-1} + \varepsilon_t$$

(1)

where $\varepsilon_t$ is iid$(0, \sigma^2)$, $D(T_B)_t = 1(t = T_B + 1)$, 1(·) being the indicator function, and $T_B$ is the true break date, should there be a break. For a series of $T$ observations, the true break fraction is denoted $\lambda' = T_B/T$. The model (1) thus represents a process that includes a unit autoregressive root, i.e. is integrated of order one, I(1), but with a level break of magnitude $\theta$. It is possible to consider further autocorrelation through incorporating in Dickey-Fuller-type regressions lagged changes, but for ease of exposition we shall concentrate on the basic case (1) and assume it is known that such augmentation is unnecessary.

The AO test of the unit root null hypothesis is based on two regression steps. First fit

$$y_t = \mu + \delta DU_t + \varepsilon_t$$

where $DU_t = 1(t > T_B)$ and $T_B$ is an assumed break date. Second, fit to the residuals $\hat{\varepsilon}_t$
\[ \hat{e}_t = \omega D(T_B)_t + \alpha \hat{e}_{t-1} + \varepsilon_t \]

The IO test is based on fitting the single regression
\[ y_t = \mu + \delta DU_t + \theta D(T_B)_t + \alpha y_{t-1} + \varepsilon_t \quad (2) \]

In either case, all possible values of \( T_B \) are considered, and the choice of a particular value, on which a unit root test based on the \( t \)-ratio associated with \( \hat{\alpha} \) is conducted, must be made. Let \( t_{\hat{\alpha}}(T_B) \) denote the \( t \)-ratio for testing \( \hat{\alpha} = 1 \), then one possibility is to choose \( T_B \) as the value for which that statistic is a minimum; that is,
\[ \hat{T}_B = \text{argmin} t_{\hat{\alpha}}(T_B) \quad (3) \]

Alternatively, the break date can be selected as the value of \( |t_{\hat{\delta}}(T_B)| \), the statistic for testing \( \hat{\delta} = 0 \) against a two-sided alternative, yielding the strongest apparent evidence for rejection of the hypothesis of no change in level, so that
\[ \hat{T}_B = \text{argmax} |t_{\hat{\delta}}(T_B)| \quad (4) \]

We generated series of 100 observations from the model (1) with \( \varepsilon_t \) independent standard normal random variables for \( \theta = 2.5, 5, \) and 10, and true break fractions \( \lambda' = 0.25, 0.5, \) and 0.75. Here and throughout, finite sample critical values, obtained by simulating (1) in the no break case, \( \theta = 0 \), were employed for the tests, which are thus by construction correctly sized in this case. We carried out both AO and IO tests, using both (3) and (4) to select a breakpoint. Empirical sizes of the four tests are given in Table 1.

### Table 1

| \( \theta \) | \( \lambda' \) | \( \text{argmax} |t_{\hat{\delta}}| \) | \( \text{argmin} t_{\hat{\alpha}} \) | \( \text{argmax} |t_{\hat{\delta}}| \) | \( \text{argmin} t_{\hat{\alpha}} \) | \( \text{argmax} |t_{\hat{\delta}}| \) |
|---|---|---|---|---|---|---|
| 2.5 | 0.25 | 5.18 | 5.75 | 5.70 | 5.71 | 6.07 |
| | 0.50 | 4.64 | 5.61 | 5.59 | 5.62 | 5.76 |
| | 0.75 | 4.75 | 4.78 | 4.82 | 4.80 | 5.92 |
| 5.0 | 0.25 | 4.20 | 10.65 | 10.49 | 10.56 | 13.49 |
| | 0.50 | 4.48 | 11.86 | 11.96 | 11.94 | 12.30 |
| | 0.75 | 4.18 | 8.48 | 8.64 | 8.59 | 13.46 |
| 10.0 | 0.25 | 2.24 | 40.34 | 40.35 | 40.30 | 57.67 |
| | 0.50 | 2.42 | 45.66 | 45.98 | 45.69 | 51.77 |
| | 0.75 | 2.44 | 36.94 | 37.84 | 37.59 | 58.00 |

*Note*: Model (2) refers to the IO regression with no trend component, Model (9) refers to the regression with a linear trend term included.

© Blackwell Publishers 2001
Here and throughout the paper, results are based on 10,000 replications, and all calculations were programmed in GAUSS. The results of Table 1 indicate that, with the exception of the AO case where the breakpoint is estimated through (4), all tests become seriously over-sized as the break magnitude $\theta$ increases, so that spurious rejections are likely to be generated. In the one exceptional case, the test becomes undersized in this case, presumably leading to some loss in power.

Since breakpoint estimation appears less problematic in general when (4) and similar statistics are employed, in the remainder of this paper we explore further this approach in the IO case for possible breaks in level.

III. Further Analysis of Innovational Outlier Tests for Level Breaks

The standard asymptotic analysis, on which the validity of the tests discussed in the previous section rests, takes the break magnitude $\theta$ in (1) as fixed with increasing sample size. Here we follow Leybourne and Newbold (2000) and Kim et al. (2000) by allowing that magnitude to be proportional to the square root of the sample size. Specifically, we set

$$\theta = \sigma kT^{1/2}$$

(5)

If the data generating process is indeed characterised by (1) and (5), application of the IO regression (3) can lead to substantial asymptotic size distortions, because the estimated break date obtained from (4) does not converge to the true break date in the limit. The asymptotic distribution of $t_\delta$ under such a scenario can be shown to be (see Appendix):

$$t_\delta \Rightarrow [F(k, \lambda, \lambda')G(k, \lambda, \lambda')]^{-1/2}E(k, \lambda, \lambda')$$

(6)

where $\Rightarrow$ denotes weak convergence in distribution, $\lambda = T_B/T$, and $F(k, \lambda, \lambda')$, $G(k, \lambda, \lambda')$ and $E(k, \lambda, \lambda')$ are functionals of standard Brownian motion, given in the Appendix.

This limit distribution can be investigated further by simulation of the mean for different values of $\lambda$. We simulated the limiting distribution (6) for the illustrative case $k = 1$, $\lambda' = 0.5$ for a range of values of $\lambda$. For a particular replication we generated a series of length 1,000 of iid $N(0,1)$ random variables, used these to construct sample moments that converge weakly to the functionals of Brownian motion implicit in $F(k, \lambda, \lambda')$, $G(k, \lambda, \lambda')$ and $E(k, \lambda, \lambda')$, and then computed an estimate of (6) using these sample moments. This procedure was repeated 10,000 times for all possible values of $\lambda$ in the range (0.3, 0.7) (i.e. steps of 0.001) and the simulated mean of the asymptotic distribution obtained for each $\lambda$. The results are graphed in

© Blackwell Publishers 2001
Figure 1, with annotation for the values of the simulated mean where $\lambda$ is close to the true break point.

The mean of the asymptotic distribution of $t_{\hat{\lambda}}$ is thus shown to be maximised at $\lambda = 0.499$, i.e. $T_B' = 1$, with a discontinuity arising at $\lambda = \lambda'$. Use of (4) for selection of the break date when the break magnitude is characterised by (5) will therefore lead, in the limit, to an estimated break date one observation before the true break. Asymptotic size distortions in the innovational outlier unit root test might then be expected, given the results of Kim et al. (2000), who demonstrate that spurious rejections can occur in such tests if the break date is placed immediately before the true date.

Although our analysis of (6) has necessarily been performed by simulation, it is possible to perform a partially analytical investigation based on asymptotic mean approximations of the functionals of Brownian motion involved in $F(k, \lambda, \lambda')$, $G(k, \lambda, \lambda')$ and $E(k, \lambda, \lambda')$. Replacing each of these functionals by their respective means yields, after some simplification, the following approximation, $E^*(t_{\hat{\lambda}})$.
\[ E^*(t_δ) \approx \begin{cases} 
 \frac{k[1 + 2k^2(1 - \lambda')][\lambda(1 - \lambda) + 2k^2\lambda(1 - \lambda')(\lambda' - \lambda)]^{1/2}}{(1 + k^2)^{1/2}[1 - \lambda + 2k^2(1 - \lambda')(\lambda' - \lambda)][1 + 2k^2\lambda'(1 - \lambda')]^{1/2}} & \lambda < \lambda' \\
 0 & \lambda = \lambda' \\
 -\frac{k(1 - \lambda)^{1/2}}{(1 + k^2)^{1/2}[\lambda + 2k^2\lambda'(\lambda - \lambda')]^{1/2}[1 + 2k^2\lambda'(1 - \lambda')]^{1/2}} & \lambda > \lambda' 
\end{cases} \]

Figure 2 provides a plot of this function for different \( \lambda \) using the same parameter values as in the simulation above (\( k = 1, \lambda' = 0.5 \)). In this example, as \( \lambda \to \lambda' \) from below, \( E^*(t_δ) \approx 1.15 \), when \( \lambda = \lambda' \), \( E^*(t_δ) \approx 0 \), and as \( \lambda \to \lambda' \) from above, \( E^*(t_δ) \approx -0.58 \). The qualitative result is again clearly highlighted, with the function of means being maximised at \( T_B' = 1 \).

The natural modification to this break date selection procedure is to simply estimate the break point at one observation later than that suggested by use of (4), i.e. modifying the selection criterion to

\[ \hat{T}_B = 1 + \arg\max|t_δ(T_B)| \]
By appealing to the result (6), we can deduce that this procedure will choose the correct break date asymptotically under the null, and should consequently overcome any asymptotic size distortions in the unit root test.¹

The scaling of the break specified by (5) is useful for modelling the test’s asymptotic behaviour in the presence of substantially sized breaks. However, a particular break of any size can also be treated as fixed, and Perron and Vogelsang (1992) derived the asymptotic distribution of the unit root test statistic $t_{\hat{\lambda}}$ under such a fixed break assumption. When the break size is indeed fixed, $t_{\hat{\lambda}}$ has the same asymptotic distribution irrespective of whether (4) or (8) is used. Perron and Vogelsang (1992) derived this distribution with an a priori assumption that the direction of the break is known (equation 10 of their paper); without that assumption, the distribution takes the same form with $\hat{\lambda}^* = \arg\min_{\lambda \in \Lambda} Z(\lambda)$ replaced by $\hat{\lambda}^* = \arg\max_{\lambda \in \Lambda} |Z(\lambda)|$.

Asymptotic critical values can be simulated using a method analogous to that of Perron and Vogelsang (1992) and are reported in the Model (2) part of Table 2. The asymptotic theory relies on an assumption that the space of values for $\lambda$ is restricted to a closed subset of (0,1); for our purposes we restricted $\lambda$ to lie between the conventionally chosen points (0.2,0.8). Finite sample critical values of the test when (8) is employed for the break date selection are also provided in Table 2. These critical values were derived by repeated application of the test to generated driftless random walks (without breaks) of appropriate length. Convergence of these finite sample critical values to their asymptotic counterparts is slow in contrast to when the

<table>
<thead>
<tr>
<th>$T$</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>-3.63</td>
<td>-3.93</td>
<td>-4.56</td>
<td>-4.12</td>
<td>-4.42</td>
<td>-5.00</td>
</tr>
<tr>
<td>100</td>
<td>-3.72</td>
<td>-4.01</td>
<td>-4.58</td>
<td>-4.17</td>
<td>-4.44</td>
<td>-4.97</td>
</tr>
<tr>
<td>150</td>
<td>-3.77</td>
<td>-4.04</td>
<td>-4.60</td>
<td>-4.20</td>
<td>-4.47</td>
<td>-4.99</td>
</tr>
<tr>
<td>200</td>
<td>-3.79</td>
<td>-4.07</td>
<td>-4.61</td>
<td>-4.22</td>
<td>-4.50</td>
<td>-4.99</td>
</tr>
<tr>
<td>500</td>
<td>-3.89</td>
<td>-4.16</td>
<td>-4.71</td>
<td>-4.31</td>
<td>-4.59</td>
<td>-5.10</td>
</tr>
<tr>
<td>$\infty$</td>
<td>-4.04</td>
<td>-4.33</td>
<td>-4.89</td>
<td>-4.45</td>
<td>-4.72</td>
<td>-5.19</td>
</tr>
</tbody>
</table>

Note: Model (2) refers to the IO regression with no trend component, Model (9) refers to the regression with a linear trend term included.

¹An alternative selection procedure was also considered, based on maximising the joint F-statistic on the dummy variable coefficients $\delta$ and $\theta$ in (2). However, the limiting distribution of this statistic was found to depend on nuisance parameters, precluding its use in application.
original selection rule (4) is employed. To confirm that the critical values are indeed converging, we conducted a simulation for $T = 5000$, resulting in critical values very close to those in the limit ($-4.03$, $-4.31$ and $-4.85$ for tests at the 10%-,$5%$- and 1%-level respectively).

It is often required to generalise the above procedures by including a linear trend component in the innovational outlier regression; that is, fitting

$$y_t = \mu + \beta t + \delta DU_t + \theta D(T_B), + \alpha y_{t-1} + \epsilon_t \quad (9)$$

Parallel results to those above apply to such a model, with application of (4) resulting in an estimated break date one observation too soon in the limit, under the null hypothesis for large breaks. Adoption of (8) in place of (4) is appropriate once again. Asymptotic critical values for the corresponding unit root test can again be derived by simulation of the null limiting distribution in the case of no break or, equivalently, a break whose magnitude is independent of $T$; this distribution is given in Vogelsang and Perron (1998, equation 11) and the critical values obtained are provided in the Model (9) part of Table 2. In addition, finite sample critical values for this test were simulated using methods as described above, and are also given in the table.

IV. Size, Power and Estimated Break Date Simulations

The theoretical results of the previous section suggest asymptotic size distortions may arise in the standard innovational outlier test when there is a break under the null, a conclusion found to be relevant for series of 100 observations, as shown by the simulation results of Table 1. Parallel results for the case where a linear trend is included, again with critical values that are by construction correct when $\theta = 0$ in (1), are also given in Table 1.

Table 3 presents the percentage of times particular break dates are chosen by each test. In line with our expectations, the standard tests have a strong tendency to select a break date one period before the true break when $\theta$ is large. It is this feature of the test that drives the size distortions observed in Table 1, and the extent to which $T_B' - 1$ is selected matches the extent to which the unit root test is over-sized. Where they are comparable, our results for Model (9) in Tables 1 and 3 are in line with those of Lee and Strazicich (1999). Adoption of the modified break date selection rule will clearly shift the estimated break dates forward by one period. The timing of large breaks is then correctly identified, overcoming the problems of the standard test.

To assess the behaviour of our modified test for a typical sample size, we repeated the simulation experiments that generated the results of Table 1, but

---

For example, simulated critical values for $T = 100$ when the break date selection rule is (4) are $-4.02$, $-4.31$ and $-4.91$ for tests at the 10%-,$5%$- and 1%-level respectively.
now basing break date selection on (8), and using the critical values of Table 2. Empirical sizes are shown in Table 4. It appears that this modified test is not seriously over-sized, though it does tend to be somewhat conservative when the break magnitude is very large.

It is also of interest to compare rejection percentages of the standard and

---

**TABLE 3**

Percentage of times particular break dates are selected when a break occurs under the null: $T = 100$.

<table>
<thead>
<tr>
<th>Panel A. Model (2), $\hat{T}<em>B = \arg\max_j t</em>\delta(T_B)$</th>
<th>$\theta = 2.5$</th>
<th>$\theta = 5.0$</th>
<th>$\theta = 10.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda'$</td>
<td>$0.25$</td>
<td>$0.50$</td>
<td>$0.75$</td>
</tr>
<tr>
<td>$&lt; T_B - 4$</td>
<td>$4.41$</td>
<td>$38.54$</td>
<td>$72.08$</td>
</tr>
<tr>
<td>$T_B - 4$</td>
<td>$2.58$</td>
<td>$1.99$</td>
<td>$2.19$</td>
</tr>
<tr>
<td>$T_B - 3$</td>
<td>$2.73$</td>
<td>$2.46$</td>
<td>$2.73$</td>
</tr>
<tr>
<td>$T_B - 2$</td>
<td>$3.29$</td>
<td>$3.36$</td>
<td>$3.68$</td>
</tr>
<tr>
<td>$T_B - 1$</td>
<td>$8.35$</td>
<td>$8.89$</td>
<td>$9.39$</td>
</tr>
<tr>
<td>$T_B'$</td>
<td>$0.56$</td>
<td>$0.56$</td>
<td>$0.72$</td>
</tr>
<tr>
<td>$T_B + 1$</td>
<td>$1.27$</td>
<td>$0.96$</td>
<td>$1.25$</td>
</tr>
<tr>
<td>$T_B + 2$</td>
<td>$0.99$</td>
<td>$0.93$</td>
<td>$1.20$</td>
</tr>
<tr>
<td>$T_B + 3$</td>
<td>$0.88$</td>
<td>$1.13$</td>
<td>$1.49$</td>
</tr>
<tr>
<td>$T_B + 4$</td>
<td>$1.09$</td>
<td>$1.05$</td>
<td>$1.68$</td>
</tr>
<tr>
<td>$&gt; T_B + 4$</td>
<td>$73.85$</td>
<td>$40.13$</td>
<td>$3.59$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Model (9), $\hat{T}<em>B = \arg\max_j t</em>\delta(T_B)$</th>
<th>$\theta = 2.5$</th>
<th>$\theta = 5.0$</th>
<th>$\theta = 10.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda'$</td>
<td>$0.25$</td>
<td>$0.50$</td>
<td>$0.75$</td>
</tr>
<tr>
<td>$&lt; T_B - 4$</td>
<td>$4.74$</td>
<td>$38.93$</td>
<td>$72.69$</td>
</tr>
<tr>
<td>$T_B - 4$</td>
<td>$2.44$</td>
<td>$1.98$</td>
<td>$1.78$</td>
</tr>
<tr>
<td>$T_B - 3$</td>
<td>$2.59$</td>
<td>$2.33$</td>
<td>$2.27$</td>
</tr>
<tr>
<td>$T_B - 2$</td>
<td>$3.52$</td>
<td>$3.08$</td>
<td>$2.96$</td>
</tr>
<tr>
<td>$T_B - 1$</td>
<td>$10.37$</td>
<td>$10.24$</td>
<td>$10.53$</td>
</tr>
<tr>
<td>$T_B$</td>
<td>$0.44$</td>
<td>$0.38$</td>
<td>$0.35$</td>
</tr>
<tr>
<td>$T_B + 1$</td>
<td>$1.10$</td>
<td>$1.15$</td>
<td>$1.08$</td>
</tr>
<tr>
<td>$T_B + 2$</td>
<td>$1.04$</td>
<td>$1.10$</td>
<td>$1.19$</td>
</tr>
<tr>
<td>$T_B + 3$</td>
<td>$1.19$</td>
<td>$1.12$</td>
<td>$1.46$</td>
</tr>
<tr>
<td>$T_B + 4$</td>
<td>$1.16$</td>
<td>$1.13$</td>
<td>$1.67$</td>
</tr>
<tr>
<td>$&gt; T_B + 4$</td>
<td>$71.41$</td>
<td>$38.56$</td>
<td>$4.02$</td>
</tr>
</tbody>
</table>

*Note:* Model (2) refers to the IO regression with no trend component, Model (9) refers to the regression with a linear trend term included.
modified tests when a break occurs under the alternative hypothesis. We simulated the model

\[ y_t = \mu + \delta DU_t + \alpha y_{t-1} + \epsilon_t \]  

(10)

with \( T = 100, \mu = 0, \delta = \gamma(1 - \alpha) \) with \( \gamma = 0, 2.5, 5, 10, \alpha = 0.7, 0.8, 0.9 \), and \( \lambda' = 0.25, 0.5, 0.75 \). The additional factor \((1 - \alpha)\) is included in \( \delta \) to ensure a mean shift in (10) of size \( \gamma \), corresponding to the null simulations. The results are given in Table 5.

The estimated rejection proportions for the standard and modified tests are not dramatically different, and neither test is uniformly more powerful than the other. The modified test has some advantages over the standard test when no break occurs under the null hypothesis, and also when the break is relatively small. For larger breaks, the ranking is reversed, although the magnitudes involved are not generally great. Rejections are also more frequent for Model (2) than Model (9) in each case as would be expected.\(^3\) It should be added that, for the standard test, rejection percentages cannot properly be identified as ‘powers’ when there is a large break, as the test is seriously over-sized in that case.

The modified test can therefore be strongly recommended. In situations where there is no break or only a small break under the null or the alternative, use of the modified break date selection rule generates a test with correct finite sample size and slightly superior power to the standard test. When the break

\(^3\)Note that the appearance of power below empirical size for some of the tests occurs because the null model is not a limiting case of the alternative as \( \alpha \) tends to one, though both are nested in testing equations such as (2).
TABLE 5
Estimated rejection percentages of nominal 5%-level unit root tests when a break occurs under the alternative (10): $T = 100$, $\delta = \gamma(1 - \alpha)$

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\lambda'$</th>
<th>$\alpha = 0.7$</th>
<th>$\alpha = 0.8$</th>
<th>$\alpha = 0.9$</th>
<th>$\alpha = 0.7$</th>
<th>$\alpha = 0.8$</th>
<th>$\alpha = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td></td>
<td>78.98</td>
<td>39.42</td>
<td>12.35</td>
<td>70.63</td>
<td>34.64</td>
<td>11.85</td>
</tr>
<tr>
<td>2.5</td>
<td>0.25</td>
<td>80.63</td>
<td>41.36</td>
<td>13.16</td>
<td>68.79</td>
<td>33.09</td>
<td>12.14</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>83.55</td>
<td>45.98</td>
<td>14.11</td>
<td>66.20</td>
<td>31.62</td>
<td>11.29</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>81.56</td>
<td>40.43</td>
<td>11.68</td>
<td>68.43</td>
<td>32.44</td>
<td>19.79</td>
</tr>
<tr>
<td>5.0</td>
<td>0.25</td>
<td>89.15</td>
<td>49.41</td>
<td>14.19</td>
<td>72.18</td>
<td>31.61</td>
<td>11.05</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>89.36</td>
<td>49.82</td>
<td>14.92</td>
<td>69.38</td>
<td>29.45</td>
<td>10.13</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>87.03</td>
<td>43.89</td>
<td>10.38</td>
<td>72.33</td>
<td>30.58</td>
<td>9.43</td>
</tr>
<tr>
<td>10.0</td>
<td>0.25</td>
<td>99.53</td>
<td>82.94</td>
<td>23.09</td>
<td>95.20</td>
<td>50.76</td>
<td>9.60</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>99.38</td>
<td>80.73</td>
<td>20.89</td>
<td>95.25</td>
<td>51.36</td>
<td>9.15</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>98.66</td>
<td>70.25</td>
<td>11.37</td>
<td>95.43</td>
<td>52.59</td>
<td>8.19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\lambda'$</th>
<th>$\alpha = 0.7$</th>
<th>$\alpha = 0.8$</th>
<th>$\alpha = 0.9$</th>
<th>$\alpha = 0.7$</th>
<th>$\alpha = 0.8$</th>
<th>$\alpha = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td></td>
<td>85.46</td>
<td>47.27</td>
<td>15.02</td>
<td>75.85</td>
<td>38.70</td>
<td>12.39</td>
</tr>
<tr>
<td>2.5</td>
<td>0.25</td>
<td>85.17</td>
<td>46.74</td>
<td>15.04</td>
<td>71.28</td>
<td>35.58</td>
<td>11.92</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>85.29</td>
<td>49.08</td>
<td>15.80</td>
<td>70.23</td>
<td>33.89</td>
<td>11.51</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>84.23</td>
<td>44.79</td>
<td>13.99</td>
<td>71.20</td>
<td>35.22</td>
<td>11.20</td>
</tr>
<tr>
<td>5.0</td>
<td>0.25</td>
<td>87.40</td>
<td>48.29</td>
<td>14.40</td>
<td>68.04</td>
<td>29.88</td>
<td>10.44</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>88.00</td>
<td>48.89</td>
<td>15.68</td>
<td>66.36</td>
<td>28.90</td>
<td>9.92</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>86.48</td>
<td>45.10</td>
<td>11.96</td>
<td>68.65</td>
<td>28.79</td>
<td>9.42</td>
</tr>
<tr>
<td>10.0</td>
<td>0.25</td>
<td>97.51</td>
<td>73.39</td>
<td>18.79</td>
<td>85.07</td>
<td>39.66</td>
<td>7.65</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>97.11</td>
<td>72.19</td>
<td>17.16</td>
<td>85.18</td>
<td>39.65</td>
<td>7.37</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>95.83</td>
<td>63.96</td>
<td>11.38</td>
<td>84.95</td>
<td>39.79</td>
<td>6.69</td>
</tr>
</tbody>
</table>

Panel A. $\hat{T}_B = \text{argmax}\{|t_\delta(T_B)|\}$

Panel B. $\hat{T}_B = 1 + \text{argmax}\{|t_\delta(T_B)|\}$

Note: Model (2) refers to the IO regression with no trend component, Model (9) refers to the regression with a linear trend term included.

considered is larger, the advantages of the modification are even greater: while the standard test is substantially over-sized, the modified test remains either correctly sized or slightly under-sized depending on the break magnitude, while losing relatively little in ‘power’ to the over-sized standard test.  

An alternative method of break date selection is to choose the date which minimises the variance of the residuals from the IO regression (2) or (9) (c.f. Nunes et al., 1997). However, we found that this approach was no better than the recommended selection rule (8), and was in some cases worse. In particular, although the correct break date was chosen for large breaks, the resultant unit root test was severely under-sized for breaks of five and ten standard deviations, and performed less favourably under the alternative hypothesis.  

© Blackwell Publishers 2001
V. Summary

We have analysed in some detail the performance of unit root tests when an endogenously determined break in level is permitted. By contrast with previous analyses, our asymptotic theory suggests the possibility that, in the case of the innovational outlier model, an approach that has been recommended and applied in earlier research can be seriously over-sized, and therefore generate spurious rejections, for practically interesting sample sizes when there is a large break under the null hypothesis. This conjecture is confirmed by simulation evidence. The asymptotic analysis also suggests that this problem might be overcome by a modest modification of the test. Simulation evidence also confirms that this is the case, and also that nothing is lost in power, in those cases where the sizes of the standard test are relatively reliable, by adopting the modified test.

A heuristic explanation of why the standard asymptotics can prove misleading might be useful. Suppose that the true generating model is a random walk, but with a single break in level. Then the series of first differences, which contains all useful information, will be white noise with a single outlier. Holding fixed the magnitude of that outlier, its location will become practically impossible to detect as the sample size grows, leading to the conclusion that, asymptotically, associated analyses will be invariant to the magnitude of the outlier, which can hence be taken as zero. Of course, for large outliers, this is far from the case in finite samples. Contrast this with the case where a break in slope of fixed magnitude is considered. Now the series of first differences of a random walk are white noise with a single break in mean. Then, as the sample size grows, location of this breakpoint becomes almost certain. This explains the conclusion that, asymptotically, associated unit root tests behave as tests when the break date is known, whatever the magnitude of any non-zero break. Naturally, however, this conclusion will not hold when the break amount is precisely zero.

Date of Receipt of Final Manuscript: September 2001.

Acknowledgment

The authors would like to thank Dick van Dijk for useful comments concerning the modification to the test procedure.

References


Appendix

Proof of (6): The regression (2) can alternatively be written as

\[ \Delta y_t = \mu + \delta DU_t + \theta D(T_B)_t + \rho y_{t-1} + \varepsilon_t \]

where \( \rho = \alpha - 1 \). The OLS estimators of the regression parameters are then given by
Straightforward but tedious algebra gives the following limits of the right hand side terms:

\[
T^{-3/2} \sum_{t=1}^{T} y_{t-1} \Rightarrow k \sigma (1 - \lambda') + \sigma \int_{0}^{1} W(r) \, dr
\]

\[
T^{-3/2} \sum_{t=B+1}^{T} y_{t-1} \Rightarrow k \sigma [1 - \lambda - I_{1}(\lambda' - \lambda)] + \sigma \int_{\lambda}^{1} W(r) \, dr
\]

\[
T^{-1/2} y_{TB} \Rightarrow (1 - I_{1})k \sigma + \sigma W(\lambda)
\]

\[
T^{-2} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow k^2 \sigma^2 (1 - \lambda') + \sigma^2 \int_{0}^{1} W(r)^2 \, dr + 2k \sigma^2 \int_{\lambda'}^{1} W(r) \, dr
\]

\[
T^{-1/2} \sum_{t=1}^{T} \Delta y_{t} \Rightarrow k \sigma + \sigma W(1)
\]

\[
T^{-1/2} \sum_{t=B+1}^{T} \Delta y_{t} \Rightarrow I_{1}k \sigma + \sigma [W(1) - W(\lambda)]
\]

\[
T^{-1/2} \Delta y_{TB+1} \Rightarrow p k \sigma, \ T_B = T'_{B} \quad \Delta y_{TB+1} = \varepsilon_{TB+1}, \ T_B \neq T'_{B}
\]

\[
T^{-1} \sum_{t=1}^{T} \Delta y_{t} y_{t-1} \Rightarrow k \sigma^2 W(1) + \frac{\sigma^2}{2} [W(1)^2 - 1]
\]

where \( I_{1} = 1(T_B \leq T'_{B}) \) and \( W(r) \) denotes a standard Wiener process. The limiting distributions of the (scaled) regression parameter estimators can then be derived:
where $I_2 = 1(T_B \neq T_B')$,

$$
A = \begin{bmatrix}
1 & 1 - \lambda & 1 - I_2 & \sigma C_1 \\
1 - \lambda & 1 - \lambda & 1 - I_2 & \sigma C_2 \\
0 & 0 & 1 & 0 \\
\sigma C_1 & \sigma C_2 & (1 - I_2)\sigma W(\lambda) & \sigma^2 C_3
\end{bmatrix},
$$

$$
B = \begin{bmatrix}
\sigma C_4 \\
\sigma C_5 \\
(1 - I_2)k\sigma + I_2\varepsilon_{T_1+1} \\
\sigma^2 C_6
\end{bmatrix},
$$

$$
C_1 = k(1 - \lambda') + \int_0^1 W(r).dr,
$$

$$
C_2 = k[1 - \lambda - I_1(\lambda' - \lambda)] + \int_{\lambda'}^1 W(r).dr,
$$

$$
C_3 = k^2(1 - \lambda') + \int_0^1 W(r)^2.dr + 2k\int_{\lambda'}^1 W(r).dr,
$$

$$
C_4 = k + W(1),
$$

$$
C_5 = I_1 k + W(1) - W(\lambda),
$$

$$
C_6 = kW(1) + \frac{1}{2}[W(1)^2 - 1]
$$

The estimated standard deviation of $\hat{\delta}$, $\hat{\sigma}$, has limiting distribution

$$
T\hat{\sigma}_\delta^2 \Rightarrow s^2 A_{2,2}^{-1}
$$

where $s^2$ is the standard estimator of $\sigma^2$ and $A_{2,2}^{-1}$ denotes the (2, 2) element of $A^{-1}$. Now, $s^2$ has probability limit

$$
s^2 \xrightarrow{p} (1 + I_2 k^2)\sigma^2
$$
The limiting distribution of the $t$-ratio on $\hat{\delta}$, $t_{\hat{\delta}} = \hat{\delta}/\hat{\sigma}_{\hat{\delta}}$ can therefore be written as

$$t_{\hat{\delta}} \Rightarrow (A^{-1}B)_{2}[(1 + I_{2}k^{2})\sigma^{2}A_{2,2}^{-1}]^{-1/2}$$

where $(A^{-1}B)_{2}$ denotes the second element of $A^{-1}B$. Substituting in for $(A^{-1}B)_{2}$ and $A_{2,2}^{-1}$ provides the full expression for the asymptotic distribution of $t_{\hat{\delta}}$:

$$t_{\hat{\delta}} \Rightarrow [F(k, \lambda, \lambda')G(k, \lambda, \lambda')]^{-1/2}E(k, \lambda, \lambda')$$

where

$$E(k, \lambda, \lambda') = C_{4}[C_{1}C_{2} - (1 - \lambda)C_{3}] + C_{5}(C_{3} - C_{1}^{2}) + C_{6}[(1 - \lambda)C_{1} - C_{2}] + (1 - I_{2})k[W(\lambda)C_{2} + C_{1}^{2} - (1 - \lambda)W(\lambda)C_{1} - \lambda C_{3} - C_{1}C_{2}],$$

$$F(k, \lambda, \lambda') = (1 + I_{2}k^{2})(C_{3} - C_{1}^{2}),$$

$$G(k, \lambda, \lambda') = \lambda(1 - \lambda)C_{3} + 2(1 - \lambda)C_{1}C_{2} - C_{2}^{2} - (1 - \lambda)C_{1}^{2}$$