Unit Root Tests in a Time Series Regression with Structural Changes

K. Maekawa[a], Z. He[b] and K. Tee[c]

[a] *Faculty of Economics, Hiroshima University (maekawa@hiroshima-u.ac.jp)*
[b] *Faculty of Business, Hiroshima Aki Women's University (kaa@hiroshima-aki-u.ac.jp)*
[c] *Faculty of Economics, Hiroshima University (khtee@hiroshima-u.ac.jp)*

Abstract: A number of tests for unit roots were proposed in the presence of structural changes in I(1) and I(0) model when the numbers of break points are or are not known (though their locations are unknown). Finite sample simulations often showed that their methods resulted in remarkable negative biases in the break points estimates. Our paper attempts to eliminate the negative biases by utilizing the weighted symmetric estimation.

***Keywords*:** Unit root test; Structural change; Weighted symmetric estimation

1. INTRODUCTION

It is commonly said among macroeconomists that Japanese GDP seems to have structural breaks twice between the mid- 50's and mid- 90's. In fact visual inspection of Figure 1 of the GDP graph in Hatanaka and Yamada [1999, abbreviated as HY hereafter] looks as if there are two break points: the first one is about 1971, and the second about 1991. But it is not easy to determine the precise break points if there is a unit root in the GDP. HY attacked this problem and showed the two break points in Japanese GDP over the period of 1957-1995 by using their proposed statistical method. They proposed unit root tests in the presence of structural changes in I(1) and I(0) models when the number of break points are or are not known. Their method consists of two steps: estimating break points and testing a unit root. However their simulation study showed that the estimated break points in the first step had serious downward bias. We focus our attention on eliminating this downward bias and confine ourselves to the case when there are two break points though their locations are unknown. This paper is a shorter version of our working paper, Maekawa, He, and Tee (2001), where derivations and proofs are given (available on request).

2. REVIEW of HY

To deal with structural changes the following notation and dummy variables are introduced. Let λ_1 and λ_2 with $\lambda_1 > \lambda_2$ be real numbers in (0,1). HY assumed that there are two break points at time $T_i = [T\lambda_i], i = 1,2$ in the sample period, $1,2,\cdots,T$, where [ ] denotes the integer part of the argument its and $\lambda = (\lambda_1, \lambda_2)$ is called (the vector of) break fractions. Define

$$dt(t,\lambda_i) = \begin{cases} 0 & \text{if } 1 \le t \le T_i, \\ t - T_i & \text{otherwise.} \end{cases}, i = 1,2. \quad (1)$$

By using these notations, HY's model can be written as

$$y_t = \beta_1 + \beta_2 t + \gamma_1 dt(t,\lambda_1) + \gamma_2 dt(t,\lambda_2) + u. \quad (2)$$

It is assumed that the disturbance u_t is a pth-order autoregressive (AR(p)) process,

$$(1 - a_1 L - a_2 L^p - \cdots - a_p L^p) u_t = \varepsilon_t,$$

where L is the lag operator and $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$. Let

$$A(L) = (1 - a_1 L - a_2 L^2 - \cdots - a_p L^p),$$

then as is well-known $A(L)$ can be rewritten as

$$A(L) = 1 - (a_1 L + a_2 L^2 + \cdots + a_p L^p) \quad (3)$$
$$= B(L)(1 - L) - \alpha L,$$

where

$$B(L) = 1 - (b_1 L + b_2 L^2 + \cdots + b_{p-1} L^{p-1}),$$
$$b_j = -(a_{j+1} + \cdots + a_p), j = 1,2,\cdots, p-1,$$
$$\alpha = -[1 - (a_1 + a_2 + \cdots + a_p)].$$

By letting $du(t, \lambda_i) = \Delta dt(t, \lambda_i)$ we have

$$du(t,\lambda_i) = \begin{cases} 0 & \text{if } 1 \le t \le T_i, \\ 1 & \text{otherwise.} \end{cases} \quad i = 1,2, \quad (4)$$

and the model (2) can be rewritten as

$$\Delta y_t \approx \mu_0 + \mu_1 du(t, \lambda_1) + \mu_2 du(t, \lambda_2) + \delta_0 t$$
$$+ \delta_1 dt(t, \lambda_1) + \delta_2 dt(t, \lambda_2) + \alpha y_{t-1} + b_1 \Delta y_{t-1} \quad (5)$$
$$+ \cdots + b_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

where all the parameters $\mu_i, \delta_i, i = 0,1,2$, can be expressed in terms of $\beta_1, \beta_2, \gamma_1, \gamma_2, a_1, \cdots, a_p$ (see Eq.(7) in p.260 in HY). This is called the Dickey-Fuller equation.

HY considered testing the null hypothesis H_0: $\{u_t\}$ is I(1), i.e., $\alpha = \delta_0 = \delta_1 = \delta_2 = 0$, against the alternative hypothesis H_1: $\{u_t\}$ is I(0) for unknown break points (λ_1, λ_2). HY utilized a t-statistic to test $\alpha = 0$ and an F-statistic to test $\alpha = \delta_0 = \delta_1 = \delta_2 = 0$. Their procedure to detect the break points and test for a unit root is as follows: (i) estimate the position of the break points by choosing the values of λ_1 and λ_2 (T_1 and T_2) that minimize the sum of squares of Ordinary Least Squares (OLS) residuals in the Dickey-Fuller equations, (ii) substitute the estimated break points into equations and calculate the t-statistic for α, and (iii) accept or reject H_0 according to the computed t-value.

HY's Monte Carlo simulation is designed to produce data that mimics Japanese GDP over the period (1957-1995). The first break point $\bar{T}_1$=60 and second break point $\bar{T}_2$=135, while the sample size T=150. The parameter values of (2) are $\bar{\beta}_1 = 1.72, \bar{\beta}_2 = 0.021, \bar{\gamma}_1 = -0.014, \bar{\gamma}_2 = -0.002$. The disturbance $_t$ generated by $(1-0.5L)(1-0.2L)\Delta u_t = \varepsilon_t$.

HY's simulation result is given in Table 1 and Figure 1. Figure 1 shows that $\bar{T}_1$ is estimated with a negative bias, that is, the break is estimated to occur before actually does, and $\bar{T}_2$ is estimated very badly. The result clearly shows that their

Table1. Distribution of $(\hat{T}_1, \hat{T}_2)$ computed by HY.

		2nd break point			
		110-118	119-129	130-140	total
1st break point	35-43	4	10	10	24
	44-54	75	62	51	188
	55-59	98	61	53	212
	60	222	180	170	572
	61-65	1	2	0	3
	66-76	1	0	0	1
	77-85	0	0	0	0
	total	401	315	284	1000

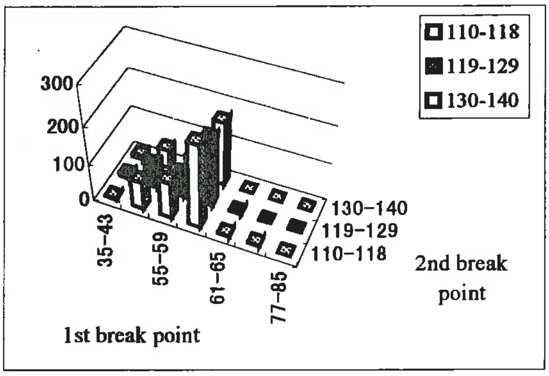

Figure 1. Distribution of $(\hat{T}_1, \hat{T}_2)$ computed by HY.

method has remarkable negative biases in the break point estimates. This paper attempts to eliminate the negative biases by utilizing the weighted symmetric (abbreviated as WS) estimation. (Park and Fuller [1994]). They points out that WS method generally has smaller mean square error than the OLS, particularly when one root is close to one in absolute value and for the model with an estimated intercept, the one-sided WS least squares test is the most powerful test.

3. WEIGHTED SYMMETRIC ESTIMATION AND THE t-TEST

First, for any given real numbers $\lambda_i, i = 1,2$, we estimate the model (2) by OLS and calculate the OLS residuals $\hat{u}_t$ to satisfy:

$$y_t = \hat{\beta}_1 + \hat{\beta}_2 t + \hat{\gamma}_1 dt(t, \lambda_1) + \hat{\gamma}_2 dt(t, \lambda_2) + \hat{u}_t \quad (6)$$
$$t = 1,2,\cdots T$$

where ^ denotes an OLS estimate.

Second, construct the "backward and forward" Dickey-Fuller equations based on the residuals $\hat{u}_t$ and the relation (3), namely,

$$\Delta \hat{u}_t = \alpha \hat{u}_{t-1} + \sum_{j=1}^{p-1} b_j \Delta \hat{u}_{t-j} + e_t, \qquad (7)$$

and

$$\Delta^{-1} \hat{u}_t = \alpha \hat{u}_{t+1} + \sum_{j=1}^{p-1} b_j \Delta^{-1} \hat{u}_{t-j} + e_t^*, \qquad (8)$$

where $\Delta^{-1} = 1 - L^{-1}, L^{-1} x_t = x_{t+1}$. We define the following weighted sum of squares of the new residuals e_t^* and e_t

$$Q_w = \sum_{t=p+1}^{T} w_t e_t^2 + \sum_{t=1}^{T-p} (1 - w_{t+1}) e_t^{*2} \qquad (9)$$

where

$$w_t = \begin{cases} 0 & t = 1,2,\cdots p; \\ \dfrac{t-p}{T-2p+2} & t = p+1, p+2, \cdots, T-p+1; \\ 1 & t = T-p+2, T-p+3, \cdots, T. \end{cases}$$

The WS estimator of $\Gamma = (\alpha, b_1, \cdots, b_{p-1})'$ is obtained by minimizing Q_w, and the resulting estimator is denoted by $\hat{\Gamma}_{ws} = (\hat{\alpha}, \hat{b}_1, \cdots, \hat{b}_{p-1})'$. Then we estimate σ_ε^2 by $s^2(\lambda) = \frac{1}{T-p} Q_w(\hat{\Gamma}_{ws})$.

Finally, for any given vector of break fractions $\lambda^{(r)}, r = 1,2,\cdots$, repeat the first and the second steps so that we obtain $s^2(\lambda^{(r)}), r = 1,2,\cdots$, and then choose $\hat{\lambda}$ satisfying

$$s^2(\hat{\lambda}) = \inf_{\lambda \in (0,1)\times(0,1)} s^2(\lambda) \qquad (10)$$

Given $\hat{\lambda}$, we can construct the t-statistic for testing $\alpha = 0$, using residuals based on the WS estimator.

Our t-statistic is defined by

$$\tilde{\tau}_{ws,\alpha} = s^2(\hat{\lambda})^{-\frac{1}{2}} \left\{ x'_{-1} \left[I - P(X_1(\hat{\lambda})) \right] x_{-1} \right\}^{-\frac{1}{2}} \times \left\{ x'_{-1} \left[I - P(X_1(\hat{\lambda})) \right] \Delta x \right\} \qquad (11)$$

where the detailed definitions of notations, derivations and the result are omitted here but are given in our working paper, Maekawa, He, and Tee (2001). We derived the nonstandard asymptotic distribution of $\tilde{\tau}_{ws,\alpha}$.

4. FINITE SAMPLE SIMULATION

We carried out a Monte Carlo study using the data generating process as in HY, that is $\bar{T}_1$=60, $\bar{T}_2$=135, while T=150. The parameter of values of (2) is $\bar{\beta}_1 = 1.72, \bar{\beta}_2 = 0.021, \bar{\gamma}_1 = -0.014, \bar{\gamma}_2 = -0.002$.

The null model $\{\Delta u_t\}$ is also generated by $(1-0.5L)(1-0.2L)\Delta u_t = \varepsilon_t$. To save computation time, we also limited the interval for searching for the break points T_1 to [35,85] and T_2 to [110,140].

The finite sample distribution of $(\hat{T}_1, \hat{T}_2)$ is given in Table 2, and the t-statistic is given in Table 3. With 1,000 replications, we obtained $\hat{T}_1 = \bar{T}_1$ 953 times and $\hat{T}_2 = \bar{T}_2$ only 78 times. From Figure 2 we can see that our method can estimate $\bar{T}_1$ fairly well without any negative bias but $\bar{T}_2$ is estimated with some negative bias. It seems difficult to estimate $\bar{T}_2$ precisely because $\bar{T}_2 = 135$ is too close to the end point of the sample period and the coefficient $\bar{\gamma}_2 = -0.002$ is too small.

Table2. Distribution of $(\hat{T}_1, \hat{T}_2)$ using WS.

		2nd break point			
		110-118	119-129	130-140	total
1st break point	35-43	1	0	0	1
	44-54	6	0	1	7
	55-59	4	5	10	19
	60	364	186	403	953
	61-65	4	4	9	17
	66-76	1	2	0	3
	77-85	0	0	0	0
	total	380	197	423	1000

Table 3. Finite Sample Distributions of $\hat{t}_\alpha$ using WS.

p	0.05	0.1	0.2	0.3	0.4
$\hat{t}_\alpha$	-4.89	-4.53	-4.08	-3.81	-3.63
p	0.5	0.6	0.7	0.8	0.9
$\hat{t}_\alpha$	-3.44	-3.25	-2.98	-2.74	-2.48

p: probability that $\hat{t}_\alpha$ is greater than entry.

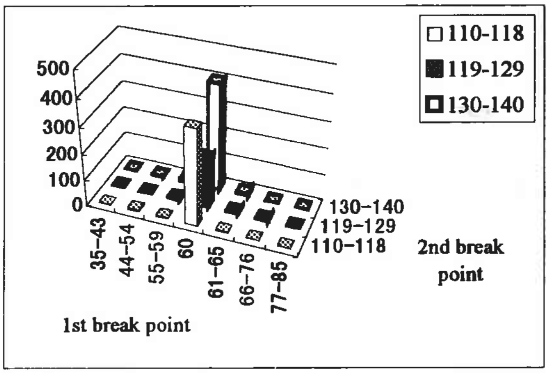

Figure 2. Distribution of $(\hat{T}_1, \hat{T}_2)$ using WS.

To determine the effects of these parameters on break point estimation, we considered two additional cases: (1) $\bar{T}_2$=120 and (2) $\bar{\gamma}_2 = -0.008$ (all other parameters remains unchanged). Our simulation shows that the HY estimator could not eliminate the negative bias in estimating $\bar{T}_2$ in the both cases. On the other hand, our WS method can consistently estimate $\hat{T}_2 = \bar{T}_2$ only 60 times out of 1000 replications in the first case ($\bar{T}_2$=120), but 602 times in the second case ($\bar{\gamma}_2 = -0.008$). The results are given in Tables 4 and 5.

Table 4. Distribution of $(\hat{T}_1, \hat{T}_2)$ using WS ($\bar{T}_2$=120).

		2nd break point			
		110-118	119-129	130-140	Total
1st break point	35-43	0	0	0	0
	44-54	4	2	0	6
	55-59	11	7	6	24
	60	417	280	253	950
	61-65	10	5	2	17
	66-76	2	1	0	3
	77-85	0	0	0	0
	Total	444	295	261	1000

Table 5. Distribution of $(\hat{T}_1, \hat{T}_2)$ using WS ($\bar{\gamma}_2 = -0.008$).

		2nd break point			
		110-118	119-129	130-140	Total
1st break point	35-43	0	0	0	0
	44-54	1	0	4	5
	55-59	2	2	14	18
	60	40	94	820	954
	61-65	0	1	11	12
	66-76	0	0	11	11
	77-85	0	0	0	0
	Total	43	97	860	1000

Comparing Figures 3 and 4, it is obvious that the performance of break point estimation in the second case of $\bar{\gamma}_2 = -0.008$ is better than in the first case of $\bar{T}_2$=120. So we can say that the cause of poor performance in estimating $\bar{T}_2$ in our initial simulation is due to the coefficient of second break point being too small.

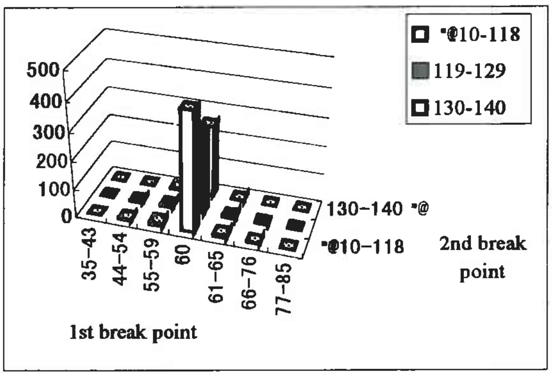

Figure 3. Distribution of $(\hat{T}_1, \hat{T}_2)$ using WS ($\bar{T}_2$=120).

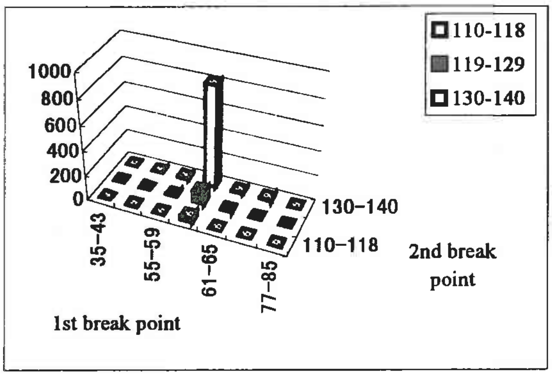

Figure 4. Distribution of $(\hat{T}_1, \hat{T}_2)$ with WS ($\bar{\gamma}_2 = -0.008$).

5. CONCLUSIONS

In this paper, we attempt to eliminate the negative biases in the break points estimates in HY's estimator by utilizing the Weighted Symmetric Estimation.

Our Monte Carlo study shows that our method can eliminate the negative bias in estimating the first break point $\bar{T}_1$, but the bias in estimation of $\bar{T}_2$ cannot be eliminated under our setup. Our Monte Carlo experiments indicate that the reason for this poor performance in estimating $\bar{T}_2$ is not because $\bar{T}_2$=135 is close to the end of the sample period, but because the coefficient of the second break point $\bar{\gamma}_2 = -0.002$ is too small.

We can expect that a unit root test based on our break point estimation would be better than a test based on biased break point estimation. Analytical and numerical comparisons of the performance of the unit root tests are now underway.

6. ACKNOWLEDGEMENT

The authors are grateful to M. Hatanaka, K. Yamada and C. McKenzie for their valuable comments.

7. REFERENCES

Hatanaka, M. and K. Yamada, A Unit Root Test in the Presence of Structural Changes in I(1) and I(0) Models, *In Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger edited by R. F. Engle and Halbert White*, Oxford University Press, 256-282, 1999.

Heon, J. P. and W. A. Fuller, Alternative Estimators and Unit Root Tests for the Autoregressive Process, *Journal of Time Series Analysis*, 415-429, 1994.

Maekawa, K., Z. He, and K. Tee, Unit Root Tests in a Time Series Regression with Structural Changes, *Working Paper of Faculty of Economics, Hiroshima University*, 2001.