



Asymptotic distribution of the sup-Wald statistic under specification errors

Terence Tai-leung Chong *

Department of Economics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Accepted 31 August 1999

Abstract

Consider a simple structural break model where $y_t = \alpha_1 + \beta_1 f(x_t) + u_t$ for $t \leq k_0$ and $y_t = \alpha_2 + \beta_2 f(x_t) + u_t$ for $t > k_0$. The timing of break and the structural parameters are unknown. Suppose the true functional form of the regressor $f(\cdot)$ is misspecified as $g(\cdot)$. We do not place too many restrictions on the functional forms of $f(\cdot)$ and $g(\cdot)$. A frequently encountered example in economics is that the true model is measured in level, but we estimate a log-linear model, i.e. when $f(x_t) = x_t$ and $g(x_t) = \log(x_t)$. For any $f(\cdot)$ and $g(\cdot)$, we derive a nonstandard limiting null distribution of the sup-Wald test statistic under some very general regularity conditions. Monte Carlo simulations support our findings. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Structural break; Functional forms; Specification errors; Brownian motions

1. Introduction

The first generation studies of structural-break models derive tests for the existence of break under the assumption that the timing of break is known a priori (Chow, 1960). The exogeneity of the break point has long been criticized, but the Chow test dominated for decades, as other tests endogenizing of the timing of break did not provide a useful distributional theory for hypothesis testing purposes (Quandt, 1960; Brown et al., 1975). The asymptotic distributions of the sup-Wald, sup LM and sup LR test statistics for structural break at unknown timing were recently puzzled out by Andrews (1993). These distributions are constructed based

* Fax: + 852-2603-5805.

E-mail address: b792703@mailserv.cuhk.edu.hk (T. Tai-leung Chong)

on correct model specification. In many instances, however, we may not know what variables should be included and in what functional form they appear in the model. Misspecification of structural-break models has been considered by Chong (1995a,b). He shows that the break point can still be consistently estimated even if the number of breaks is underspecified.

This paper considers the asymptotic null distribution of the Wald type test for parameter instability when some or all of the variables in a structural break model are misspecified. The results given here cover most types of specification errors in the independent variables. The paper is organized as follows: Section 2 presents the model and states the necessary assumptions. Section 3 derives a nonstandard asymptotic null distribution of the sup-Wald test statistic for parameter instability under specification errors. Simulations are performed in Section 4 to support our theory, and a conclusion is drawn in Section 5. All proofs are collected in an Appendix.

Before proceeding to the next section, we present some frequently used mathematical notations. Let $\|w\| = \text{tr}(w'w)^{1/2}$ be the Euclidean norm of the matrix w . $[x]$ signifies the greatest integer $\leq x$. Define a matrix to be $O_p(1)$ if all of its elements are $O_p(1)$. The symbol ' \xrightarrow{p} ' represents convergence in probability, ' \xrightarrow{d} ' represents convergence in distribution, and ' \Rightarrow ' denotes the weak convergence in $D[0,1]$ (see Billingsley, 1968 and Pollard, 1984). All limits are as the sample size $T \rightarrow \infty$ unless otherwise stated.

2. The model and assumptions

We consider a general multivariate structural-break model with a single break and i.i.d. innovations. Suppose the true model is:

$$Y = I_1 F \beta_1 + I_2 F \beta_2 + U. \quad (2.1)$$

However, we misspecify the functional form of the regressors and estimate the following model:

$$Y = I_a G \hat{\beta}_{1[\tau T]} + I_b G \hat{\beta}_{2[\tau T]} + \hat{U}, \quad (2.2)$$

where

Y is a T by 1 matrix with elements y_t , $t = 1, 2, \dots, T$.

F is a T by P matrix with the $(t, p)^{th}$ element $f_p(x_{tp})$, where $f_p(\cdot)$ is a real value function, $p = 1, 2, \dots, P$.

G is a T by P matrix with the $(t, p)^{th}$ element $g_p(x_{tp})$, where $g_p(\cdot)$ is a real value function, $p = 1, 2, \dots, P$.

U is a T by 1 matrix with the t^{th} element u_t , an i.i.d. process with zero mean and a finite variance σ^2 .

I_a is a T by T diagonal matrix with the $(t, t)^{th}$ element $1_{\{t \leq k\}}$. $1_{\{\cdot\}}$ is an indicator function that equals 1 when the statement inside the bracket is true and equals 0 otherwise.

$I_b = I - I_a$.

I_1 is a T by T diagonal matrix with the $(t, t)^{th}$ element $1\{t \leq k_0\}$.

$I_2 = I - I_1$.

$\beta_1 = (\beta_{11}, \beta_{21}, \dots, \beta_{p1})'$, $\beta_2 = (\beta_{12}, \beta_{22}, \dots, \beta_{p2})'$ are P by 1 vectors of true pre- and post-shift structural parameters, respectively.

$\hat{\beta}_{1[\tau T]}$ and $\hat{\beta}_{2[\tau T]}$ are P by 1 vectors of regression coefficients' estimates for $0 < t \leq k$ and $k < t \leq T$, respectively.

\hat{U} is a T by 1 matrix of the residuals for the misspecified model.

Let $k = [\tau T]$, where $\tau \in [0, 1]$ the break fraction.

Let Q_{ff} and Q_{gg} be positive definite and non-stochastic matrices with the $(i, j)^{th}$ element $\lim_{T \rightarrow \infty} 1/T \sum_{t=1}^T E(f_i(x_{it})f_j(x_{it}))$ and $\lim_{T \rightarrow \infty} 1/T \sum_{t=1}^T E(g_i(x_{it})g_j(x_{it}))$, respectively, Q_{fg} a non-stochastic matrix with the $(i, j)^{th}$ element $\lim_{T \rightarrow \infty} 1/T \sum_{t=1}^T E(g_i(x_{it})f_j(x_{it}))$.

Let $B_{fu}(\tau)$ and $B_{gu}(\tau)$ be P by 1 vectors of Brownian motions, and $B_{gg}(\tau)$ and $B_{fg}(\tau)$ P by P matrices of Brownian motions.

We assume that:

$$\frac{1}{T} F' I_a F \xrightarrow{p} \tau Q_{ff} \quad \text{uniformly for } \tau \in [0, 1], \quad (A1)$$

$$\frac{1}{T} G' I_a G \xrightarrow{p} \tau Q_{gg} \quad \text{uniformly for } \tau \in [0, 1], \quad (A2)$$

$$\frac{1}{T} G' I_a F \xrightarrow{p} \tau Q_{fg} \quad \text{uniformly for } \tau \in [0, 1], \quad (A3)$$

$$\frac{1}{\sqrt{T}} F' I_a U \Rightarrow B_{fu}(\tau), \quad (A4)$$

$$\frac{1}{\sqrt{T}} G' I_a U \Rightarrow B_{gu}(\tau), \quad (A5)$$

$$\sqrt{T} \left(\frac{1}{T} G' I_a G - \tau Q_{gg} \right) \Rightarrow B_{gg}(\tau), \quad (A6)$$

$$\sqrt{T} \left(\frac{1}{T} G' I_a F - \tau Q_{fg} \right) \Rightarrow B_{fg}(\tau), \quad (A7)$$

For all $1 \leq h \leq j < T$, $1 \leq p \leq P$, \exists some real numbers $r > 2$, $C_p > 0$ and $D_p > 0$ such that

$$E \left| \sum_{t=h}^j f_p(x_{tp}) u_t \right|^r \leq C_p (j-h)^{r/2} \quad (A8)$$

and

$$E \left| \sum_{t=h}^j g_p(x_{tp}) u_t \right|^r \leq D_p (j-h)^{r/2}, \quad (A9)$$

$$\inf_{0 \leq k_r < k_s \leq T} \det F'(I(k_s) - I(k_r)) F > 0, \quad (A9)$$

$$\inf_{0 \leq k_r < k_s \leq T} \det G'(I(k_s) - I(k_r)) G > 0, \quad (A10)$$

$$\theta_0 = (\tau_0, \beta'_1, \beta'_2) \in \Theta = [\underline{\tau}, \bar{\tau}] \times B \times B \subset (0,1) \times R^{2P}. \quad (A11)$$

Assumptions 1–3 imply the regressors are stationary and ergodic, trending regressors are not allowed under these assumptions. Assumptions 4–7 are for statistical inference purposes, they also bound the variation of the stochastic insignificant terms and help to construct the uniform convergence result. Assumptions 6 and 7 allow the explanatory variables in G and F to be stochastic. If all the explanatory variables in G and F are non-stochastic, then we have $B_{fg}(\tau) = B_{gg}(\tau) = 0$. Non-stochastic F and G will not affect our result. In fact, it simplifies our analysis. Assumptions 9 and 10 ensure the invertibility of the matrices defined there, so that the structural estimators are properly defined asymptotically as well as in the finite sample. Assumption 11 requires the true break point to be in a compact set in $(0,1)$, this assumption is necessary because the structural estimates $\hat{\beta}_{1[\tau T]}$ and $\hat{\beta}_{2[\tau T]}$ are not defined at the boundary of time domain.

For any given values of τ , the least-squares estimators of the pre- and post-shift parameters of (2.2) are, respectively:

$$\hat{\beta}_{1[\tau T]} = (G' I_a G)^{-1} G' I_a Y, \quad (2.3)$$

$$\hat{\beta}_{2[\tau T]} = (G' I_b G)^{-1} G' I_b Y.$$

We define the break-point estimator as the timing where the residual sum of squares is minimized, namely:

$$\hat{\tau}_T = \underset{\tau \in [\underline{\tau}, \bar{\tau}]}{\text{Arg min}} RSS_T(\tau), \quad (2.4)$$

where

$$RSS_T(\tau) = \|Y - I_a G \hat{\beta}_{1[\tau T]} - I_b G \hat{\beta}_{2[\tau T]}\|^2$$

is the residual sum of squares in the least squares estimation.

For ease of exposition, we consider here the case of a single regressor, i.e. $P = 1$. Suppose the true model is:

$$y_t = \beta_1 f(x_t) + u_t \quad \text{for } t = 1, 2, \dots, k_0. \quad (2.5)$$

$$y_t = \beta_2 f(x_t) + u_t \quad \text{for } t = k_0 + 1, \dots, T.$$

For any given τ , our estimated model is:

$$\hat{y}_t = \hat{\beta}_{1[\tau T]} g(x_t) \quad \text{for } t \leq [\tau T], \quad (2.6)$$

$$\hat{y}_t = \hat{\beta}_{2[\tau T]} g(x_t) \quad \text{for } t > [\tau T].$$

Assumptions A1–A7 imply the following:

$$S_{ff}(\tau) \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^{[\tau T]} f(x_t)^2 \xrightarrow{P} \tau \sigma_f^2 \quad \text{uniformly for } \tau \in [0, 1], \quad (A1')$$

$$S_{gg}(\tau) \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^{[\tau T]} g(x_t)^2 \xrightarrow{P} \tau \sigma_g^2 \quad \text{uniformly for } \tau \in [0, 1], \quad (A2')$$

$$S_{fg}(\tau) \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^{[\tau T]} f(x_t)g(x_t) \xrightarrow{p} \tau \sigma_{fg} \quad \text{uniformly for } \tau \in [0,1], \quad (\text{A3}')$$

$$S_{fu}(\tau) \stackrel{\text{def}}{=} \frac{1}{\sqrt{T}} \sum_{t=1}^{[\tau T]} f(x_t)u_t \Rightarrow B_{fu}(\tau), \quad (\text{A4}')$$

$$S_{gu}(\tau) \stackrel{\text{def}}{=} \frac{1}{\sqrt{T}} \sum_{t=1}^{[\tau T]} g(x_t)u_t \Rightarrow B_{gu}(\tau), \quad (\text{A5}')$$

$$\sqrt{T}(S_{gg}(\tau) - \tau \sigma_g^2) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[\tau T]} (g(x_t)^2 - \sigma_g^2) \Rightarrow B_{gg}(\tau), \quad (\text{A6}')$$

$$\sqrt{T}(S_{fg}(\tau) - \tau \sigma_{fg}) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[\tau T]} (f(x_t)g(x_t) - \sigma_{fg}) \Rightarrow B_{fg}(\tau), \quad (\text{A7}')$$

where $B_{fu}(\tau)$, $B_{fu}(\tau)$, $B_{gg}(\tau)$ and $B_{fg}(\tau)$ are Brownian motions with zero mean and variances $\tau \sigma^2 \sigma_f^2$, $\tau \sigma^2 \sigma_g^2$, $\lim_{T \rightarrow \infty} E \left(\frac{1}{T} \sum_{t=1}^{[\tau T]} (g(x_t)^2 - \sigma_g^2) \right)^2$ and $\lim_{T \rightarrow \infty} E \left(\frac{1}{T} \sum_{t=1}^{[\tau T]} (f(x_t)g(x_t) - \sigma_{fg}) \right)^2$, respectively.

The residual sum of squares is defined as:

$$RSS_T(\tau) = \sum_{t=1}^{[\tau T]} (y_t - \hat{\beta}_{1[\tau T]} g(x_t))^2 + \sum_{t=[\tau T]+1}^T (y_t - \hat{\beta}_{2[\tau T]} g(x_t))^2.$$

3. The sup-Wald test

In this section, we will investigate the validity of the sup-Wald type test statistics under specification errors. In general, the pre- and post-shift estimators will be inconsistent when misspecification occurs. However, if there is no structural break, the probability limits of the pre- and post-shift estimators are the same. While in the presence of break, their probability limits are different. Therefore in the presence of specification errors, the Wald type test that based on the magnitude of the estimated break (Hawkins, 1987) will still be a consistent test. However, because the model is misspecified, the limiting distribution of the test statistic will be affected.

Suppose (2.1) is the true model, but (2.2) is estimated, the Wald statistic for the hypothesis $H_0: \beta_1 = \beta_2$ is defined as:

$$W_T(\tau) = \frac{T\tau(1-\tau)}{RSS_T(\tau)} (\hat{\beta}_{2[\tau T]} - \hat{\beta}_{1[\tau T]})' G' G (\hat{\beta}_{2[\tau T]} - \hat{\beta}_{1[\tau T]}). \quad (3.1)$$

Theorem 1. *If assumptions A1–A10 hold, then under $H_0: \beta_1 = \beta_2 = \beta$, $T \rightarrow \infty$, we have:*

$$W_T(\tau) \Rightarrow \frac{(\tau B_A(1) - B_A(\tau))' Q_{gg}^{-1} (\tau B_A(1) - B_A(\tau))}{\tau(1-\tau)(\sigma^2 + \beta'(Q_{ff} - Q)\beta)} \quad (3.2)$$

and

$$\sup_{\tau \in S} W_T(\tau) \xrightarrow{d} \sup_{\tau \in S} \frac{(\tau B_A(1) - B_A(\tau))' Q_{gg}^{-1} (\tau B_A(1) - B_A(\tau))}{\tau(1-\tau)(\sigma^2 + \beta'(Q_{ff} - Q)\beta)} \quad (3.3)$$

where S denotes a set whose closure lies in $(0, 1)$, $Q = Q'_{fg} Q_{gg}^{-1} Q_{fg}$ and $B_A(\tau) = (B'_{fg}(\tau) - B'_{gg}(\tau) Q_{gg}^{-1} Q_{fg})\beta + B_{gu}(\tau)$ is a P -vector Brownian motions on $[0, 1]$. If all the explanatory variables in G and F are non-stochastic, then we have $B_{fg}(\tau) = B_{gg}(\tau) = 0$ and as a result $B_A(\tau) = B_{gu}(\tau)$.

If there are no specification errors, we have $B_A(\tau) = B_{fu}(\tau)$, $Q = Q_{ff}$, and $\sup_{\tau \in S} W_T(\tau) \Rightarrow \sup_{\tau \in S} \frac{\|\tau B(1) - B(\tau)\|^2}{\tau(1-\tau)}$ where $B(\tau)$ is a P -vector of independent Brownian motions. Thus the limiting distribution reduces to the supremum of the square of a standardized tie-down Bessel process of order P . If $P = 1$ and if x_t and x_s are independent $\forall t \neq s$, Theorem 1 can be reduced to:

Corollary 2. When $P = 1$, suppose assumptions A1–A10 hold, and x_t and x_s are independent $\forall t \neq s$, then under H_0 : $\beta_1 = \beta_2 = \beta$, as $T \rightarrow \infty$, we have:

$$W_T(\tau) \Rightarrow C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2) \frac{(\tau B(1) - B(\tau))^2}{\tau(1-\tau)}, \quad (3.4)$$

$$\sup_{\tau \in S} W_T(\tau) \xrightarrow{d} C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2) \sup_{\tau \in S} \frac{(\tau B(1) - B(\tau))^2}{\tau(1-\tau)},$$

where

$$C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2) = \frac{\Psi}{\sigma^2 \sigma_g^2 + \beta^2 (\sigma_f^2 \sigma_g^2 - \sigma_{fg}^2)},$$

$$\Psi = \sigma^2 \sigma_g^2 + \beta^2 \sigma_{fg}^2 \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{[T]} \text{Var}(f(x_t)g(x_t)\sigma_{fg}^{-1} - g(x_t)^2 \sigma_g^{-2}),$$

$B(\tau)$ is a Brownian motion on $[0, 1]$.

If all the explanatory variables in G and F are non-stochastic, then $\Psi = \sigma^2 \sigma_g^2$ and $C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2) = (1 + \beta^2 \sigma^{-2} (\sigma_f^2 - \sigma_g^{-2} \sigma_{fg}^2))^{-1}$.

Thus, the asymptotic null distribution of the sup-Wald statistic is the supremum of the square of a standardized tie-down Bessel process of order one multiplied by a scaling factor. Note that the scaling factor depends on the absolute values of the structural parameters as well as on the moments of the true and misspecified functions. If there is no misspecification, we have $f(x_t) = g(x_t)$, $\sigma_f^2 = \sigma_g^2 = \sigma_{fg}$, $C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2) = 1$ and

$$\sup_{\tau \in S} W_T(\tau) \xrightarrow{d} \sup_{\tau \in S} \frac{(\tau B(1) - B(\tau))^2}{\tau(1-\tau)},$$

which is the conventional null distribution of sup-Wald test statistic for testing structural break. When $f(x_t) \sim \text{nid}(0, 1)$, $g(x_t) = f(x_t)^2$, $\sigma^2 = 1$, we have:

$$C(\beta, 1, 3, 0, 1) = \frac{\sigma_g^2 + \beta^2 E(f(x_i)^6)}{\sigma_g^2 + \sigma_g^2 \beta^2} = \frac{3 + 15\beta^2}{3(1 + \beta^2)} = \frac{1 + 5\beta^2}{1 + \beta^2}.$$

4. Experimental evidence

This experiment is to verify Corollary 2.

Using the definition of $W_T(\tau)$ in (3.1), Table 1 (A–C) shows the critical value c such that:

$$\Pr\left(\sup_{\tau \in (\lambda, 1-\lambda)} W_T(\tau) > c\right) = \alpha.$$

Table 1

(A) shows the critical values in the absence of specification errors, the values are very close to those obtained in Andrews (1993); (B) and (C) display the critical values under specification errors^a

(A) $b = 1$, (No specification errors), $C(\beta, 1, 1, 1, 1) = 1$									
	$\beta = 1$ $C(1, 1, 1, 1, 1) = 1$			$\beta = 2$ $C(2, 1, 1, 1, 1) = 1$			$\beta = 3$ $C(3, 1, 1, 1, 1) = 1$		
$\lambda \backslash \alpha$	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
0.25	6.43	7.97	11.40	6.25	7.73	11.13	6.23	7.70	11.06
0.20	6.69	8.19	11.73	6.74	8.29	11.45	6.59	8.27	11.61
0.15	7.14	8.62	12.17	7.15	8.70	12.25	7.04	8.94	12.62
0.10	7.51	9.08	12.84	7.62	9.28	12.91	7.64	9.23	12.68
(B) $b = 2$, $C(\beta, 1, 3, 0, 1) = \frac{1 + 5\beta^2}{1 + \beta^2}$									
	$\beta = 1$ $C(1, 1, 3, 0, 1) = 3$			$\beta = 2$ $C(2, 1, 3, 0, 1) = 4.2$			$\beta = 3$ $C(3, 1, 3, 0, 1) = 4.6$		
$\lambda \backslash \alpha$	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
0.25	18.47	22.71	32.59	25.59	31.82	44.57	27.86	34.77	50.22
0.20	18.74	23.03	33.35	26.96	33.08	47.04	28.95	35.84	52.52
0.15	20.16	24.63	34.74	28.08	34.09	49.46	31.01	37.91	53.08
0.10	21.44	25.84	35.08	29.10	35.25	50.00	32.70	39.11	54.25
(C) $b = 3$, $C(\beta, 1, 15, 3, 1) = \frac{25 + 238\beta^2}{25 + 10\beta^2}$									
	$\beta = 1$ $C(1, 1, 15, 3, 1) = 7.51$			$\beta = 2$ $C(2, 1, 15, 3, 1) = 15.03$			$\beta = 3$ $C(3, 1, 15, 3, 1) = 18.84$		
$\lambda \backslash \alpha$	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
0.25	34.12	45.38	75.84	66.44	88.93	142.85	81.74	105.75	175.34
0.20	37.92	50.92	82.55	70.73	93.47	150.11	88.80	113.97	191.29
0.15	40.70	53.50	88.35	76.67	98.37	167.43	98.98	128.77	203.62
0.10	46.74	62.68	100.29	88.57	114.00	182.37	108.43	141.82	223.42

^a It is apparent that the values in (B) and (C) are close to the values in (A) multiplied by the corresponding factor $C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2)$. These results support Corollary 2.

True model:

$$y_t = \beta x_t + u_t \quad t = 1, 2, \dots, T.$$

Misspecified models:

$$y_t = \beta x_t^b + u_t \quad t = 1, 2, \dots, T.$$

$T = 1000$ (sample size),

$N = 1000$ (number of replications),

$x_t \sim n.i.d. (0,1)$,

$u_t \sim n.i.d. (0,1)$,

$\{x_t\}_{t=1}^T$ and $\{u_t\}_{t=1}^T$ are independent of each other.

5. Conclusion

For decades, specification error has been an important topic in econometric modellings. This paper examines how the hypothesis testing of structural-break models is affected under specification errors. We show that under the null of no structural break and under almost any kinds of data transformations on the independent variables, the sup-Wald statistic converges in distribution to the supremum of the square of a standardized tie-down Bessel process multiplied by a scaling factor. The theoretical limiting distribution of the sup-Wald statistic is confirmed by computer simulations.

Acknowledgements

I would like to thank an anonymous referee for very helpful comments. Any remaining errors are my responsibility.

Appendix A

Proof of Theorem 1. Define

$$\begin{aligned} S_{1[\tau T]} &= \sqrt{T}((G'I_b G)^{-1} G'I_b F - (G'I_a G)^{-1} G'I_a F) \\ &= \sqrt{T} \left((G'I_b G)^{-1} G'I_b F - (1 - \tau) \left(\frac{1}{T} G'I_b G \right)^{-1} Q_{fg} + (1 - \tau) \left(\frac{1}{T} G'I_b G \right)^{-1} Q_{fg} \right. \\ &\quad \left. - \tau \left(\frac{1}{T} G'I_a G \right)^{-1} Q_{fg} + \tau \left(\frac{1}{T} G'I_a G \right)^{-1} Q_{fg} - (G'I_a G)^{-1} G'I_a F \right) \\ &= \sqrt{T} \left(\left(\frac{1}{T} G'I_b G \right)^{-1} \left(\frac{1}{T} G'I_b F - (1 - \tau) Q_{fg} \right) \right) \end{aligned}$$

$$\begin{aligned}
& + \left(\frac{1}{T} G' I_b G \right)^{-1} \left(\frac{1}{T} G' I_a G - \tau \frac{1}{T} G' G \right) \left(\frac{1}{T} G' I_a G \right)^{-1} Q_{fg} \\
& - \left(\frac{1}{T} G' I_a G \right)^{-1} \left(\frac{1}{T} G' I_a F - \tau Q_{fg} \right) \\
= & \sqrt{T} \left(\left(\frac{1}{T} G' I_b G \right)^{-1} \left(\frac{1}{T} G' I_b F - (1 - \tau) Q_{fg} \right) - \left(\frac{1}{T} G' I_a G \right)^{-1} \left(\frac{1}{T} G' I_a F - \tau Q_{fg} \right) \right. \\
& \left. + \left(\frac{1}{T} G' I_b G \right)^{-1} \left(\frac{1}{T} G' I_a G - \tau Q_{gg} - \tau \left(\frac{1}{T} G' G - Q_{gg} \right) \right) \left(\frac{1}{T} G' I_a G \right)^{-1} Q_{fg} \right).
\end{aligned}$$

Using assumptions A2, A7 and A8,

$$S_{1[\tau T]} \Rightarrow \frac{1}{\tau(1-\tau)} Q_{gg}^{-1} (\tau B_{fg}(1) - B_{fg}(\tau) - (\tau B_{gg}(1) - B_{gg}(\tau)) Q_{gg}^{-1} Q_{fg}).$$

If all the explanatory variables in G and F are non-stochastic, then

$$S_{1[\tau T]} \xrightarrow{p} 0.$$

Similarly, using assumptions A2 and A5,

$$\begin{aligned}
S_{2[\tau T]} &= \left(\frac{1}{T} G' I_b G \right)^{-1} \frac{1}{\sqrt{T}} G' I_b U - \left(\frac{1}{T} G' I_a G \right)^{-1} \frac{1}{\sqrt{T}} G' I_a U \\
&\Rightarrow \frac{1}{\tau(1-\tau)} Q_{gg}^{-1} (\tau B_{gu}(1) - B_{gu}(\tau)).
\end{aligned}$$

Under H_0 of no structural break, $\beta_1 = \beta_2 = \beta$, $\frac{1}{T} RSS_T(\tau) \xrightarrow{p} \sigma^2 + \beta'(Q_{ff} - Q)\beta$,

$$\begin{aligned}
W_T(\tau) &= \frac{T\tau(1-\tau)}{RSS_T(\tau)} (\hat{\beta}_{2[\tau T]} - \hat{\beta}_{1[\tau T]})' G' G (\hat{\beta}_{2[\tau T]} - \hat{\beta}_{1[\tau T]}) \\
&= \frac{\tau(1-\tau)}{\frac{1}{T} RSS_T(\tau)} (S_{1[\tau T]}\beta + S_{2[\tau T]})' \frac{1}{T} G' G (S_{1[\tau T]}\beta + S_{2[\tau T]}) \\
&\Rightarrow \frac{(\tau B_A(1) - B_A(\tau))' Q_{gg}^{-1} (\tau B_A(1) - B_A(\tau))}{\tau(1-\tau)(\sigma^2 + \beta'(Q_{ff} - Q)\beta)}
\end{aligned}$$

where $Q = Q'_{fg} Q_{gg}^{-1} Q_{fg}$, $B_A(\tau) = (B_{fg}(\tau) - B_{gg}(\tau) Q_{gg}^{-1} Q_{fg})\beta + B_{gu}(\tau)$ is a P-vector Brownian motions. By the Continuous Mapping Theorem, we have:

$$\sup_{\tau \in \mathbb{S}} W_T(\tau) \xrightarrow{d} \sup_{\tau \in \mathbb{S}} \frac{(\tau B_A(1) - B_A(\tau))' Q_{gg}^{-1} (\tau B_A(1) - B_A(\tau))}{\tau(1-\tau)(\sigma^2 + \beta'(Q_{ff} - Q)\beta)},$$

where \mathbb{S} is a set whose closure lies in $(0, 1)$. ■

Appendix B

Proof of Corollary 2. When $P = 1$, and if x_t and x_s are independent $\forall t \neq s$, then

$$B_A(\tau) = (B_{fg}(\tau)\sigma_{fg}^{-1} - B_{gg}(\tau)\sigma_g^{-2})\sigma_{fg}\beta + B_{gu}(\tau) \stackrel{d}{=} \Psi^{1/2}B(\tau),$$

where $\Psi = \sigma^2\sigma_g^2 + \beta^2\sigma_{fg}^2 \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{[T\tau]} \text{Var}(f(x_t)g(x_t)\sigma_{fg}^{-1} - g(x_t)^2\sigma_g^{-2})$ and $B(\tau)$ Brownian motion process in $[0, 1]$.

$$\begin{aligned} \sup_{\tau \in S} W_T(\tau) &\stackrel{d}{\rightarrow} \sup_{\tau \in S} \frac{(\tau B_A(1) - B_A(\tau))^2}{\tau(1-\tau)(\sigma^2\sigma_g^2 + \beta^2(\sigma_f^2\sigma_g^2 - \sigma_{fg}^2))} \\ &\stackrel{d}{=} C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2) \sup_{\tau \in S} \frac{(\tau B(1) - B(\tau))^2}{\tau(1-\tau)}, \end{aligned}$$

where

$$C(\beta, \sigma_f^2, \sigma_g^2, \sigma_{fg}, \sigma^2) = \frac{\Psi}{\sigma^2\sigma_g^2 + \beta^2(\sigma_f^2\sigma_g^2 - \sigma_{fg}^2)}. \blacksquare$$

References

- Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.
- Billingsley, P., 1968. *Convergence of Probability Measures*. Wiley, New York.
- Brown, R.L., Durbin, J., Evans, J.M., 1975. Techniques for testing the constancy of regression relationships over time. *J. Royal Stat. Soc. Ser. B* 37, 149–192.
- Chong, Terence Tai-leung, 1995a. Partial parameter consistency in a misspecified structural change model. *Economics Lett.* 49, 351–357.
- Chong Terence Tai-leung, 1995b. *Econometrics of multiple structural changes*, Ph.D. Dissertation, University of Rochester.
- Chow, G.C., 1960. Testing for equality between sets of coefficients in two linear regressions. *Econometrica* 28, 591–605.
- Hawkins, D.L., 1987. A test for a change point in a parametric model based on a maximal Wald-type statistic. *Sankhya* 49A, 368–376.
- Pollard, D., 1984. *Convergence of Stochastic Processes*. Springer-Verlag Inc, New York.
- Quandt, R., 1960. Tests of the hypothesis that a linear regression system obeying two separate regimes. *J. Am. Stat. Assoc.* 55, 324–330.