

## Generic Consistency of the Break-Point Estimators under Specification Errors in a Multiple-Break Model

JUSHAN BAI<sup>†</sup>, HAIQIANG CHEN<sup>‡</sup>, TERENCE TAI-LEUNG CHONG<sup>\*1</sup>, AND SERAPH XIN WANG<sup>\*</sup>

<sup>†</sup> *Department of Economics, New York University*

E-mail: [jushan.bai@nyu.edu](mailto:jushan.bai@nyu.edu)

<sup>‡</sup> *Department of Economics, Cornell University*

E-mail: [hc335@cornell.edu](mailto:hc335@cornell.edu)

<sup>\*</sup> *Department of Economics, The Chinese University of Hong Kong*

E-mail: [chong2064@cuhk.edu.hk](mailto:chong2064@cuhk.edu.hk); [seraph\\_wx@hotmail.com](mailto:seraph_wx@hotmail.com)

Received: March 2006

**Summary** This paper considers the estimation of multiple-structural-break models under specification errors. A common example in economics is that the true model is measured in level, but a linear-log model is estimated. We show that, under specification errors, if there are more than one break point and if a single-break model is estimated, the estimated break point is consistent for one of the true break points. This consistency result applies to models with multiple regressors where some or all of the regressors are misspecified. Another important contribution of this paper is that we have constructed a Sup-Wald test whose limiting distribution is not affected by model misspecification. Using this robust test, we show that the break points can be estimated sequentially one at a time. Simulation evidence and an empirical application are provided.

**Keywords:** *Multiple Changes, Misspecification, Measurement Error, Consistency.*

### 1. INTRODUCTION

The end of the last century saw significant advance in the research of structural-break models<sup>1</sup>. While earlier studies focus on the single-break model, the attention has shifted to the multiple-break model in recent years. Pioneering works on multiple-break models include Chong (1995), Bai (1997), Bai and Perron (1998), and Chong (2001). The aforementioned studies propose different methods to estimate the locations and the number of breaks and establish the consistency of the break-point estimators under correct model specification. In the presence of model misspecification, Chong (2003) shows that the break-point estimator is still consistent when there is only one break. It is not known, however, whether this consistency result can be extended to the case of multiple breaks. In this paper, we extend the work of Chong (2003) to study the consistency of the break-point estimator in a multiple-break model under specification errors. This problem is important because, at least philosophically, all models are incorrect. We investigate the robustness of the estimated break point when incorrect models are employed. The main

<sup>1</sup>Corresponding author

<sup>1</sup>For the recent development of the literature, one is referred to Perron (2006).

finding is that the break points can still be consistently estimated despite incorrect model specifications.

To present the key idea of the underlying consistency argument, consider the following measurement error model of Chang and Huang (1997):

$$y_t = \beta_1 \xi_t + \epsilon_t, \quad t \leq k_1$$

$$y_t = \beta_2 \xi_t + \epsilon_t, \quad t > k_1$$

$t = 1, \dots, T$ . The variable  $\xi_t$  is not observable, but instead, we observe

$$x_t = \xi_t + \eta_t,$$

where  $\eta_t$  is the measurement error. It is assumed that  $\epsilon_t$  is independent of  $\eta_t$  and  $\xi_t$ . In standard regression models without a break, it is well known that if  $x_t$  is used as a regressor, simple least squares cannot yield consistent estimate of the slope parameters because the new error is correlated with the regressor  $x_t$ .

We can rewrite the model with new slope coefficients without the measurement error problem. First, projecting  $\xi_t$  on  $x_t$ , we have

$$\xi_t = c x_t + e_t,$$

for some constant  $c$ , where the projection residual  $e_t$  is uncorrelated with  $x_t$  by the projection argument. Thus, the original model can be rewritten as

$$y_t = (\beta_1 c) x_t + \epsilon_t + \beta_1 e_t, \quad t \leq k_1$$

$$y_t = (\beta_2 c) x_t + \epsilon_t + \beta_2 e_t, \quad t > k_1$$

The disturbances  $\epsilon_t + \beta_1 e_t$  and  $\epsilon_t + \beta_2 e_t$  are uncorrelated with the regressor  $x_t$ . Thus, this is a standard change-point problem without measurement error. Furthermore,

$$\beta_1 \neq \beta_2$$

implies

$$\beta_1 c \neq \beta_2 c.$$

Note that  $c \neq 0$  under the measurement error setup. The case of  $c = 0$  corresponds to the case of using a regressor which is unrelated to the true regressor  $\xi_t$ . Such a situation will be ruled out by our assumptions. Since we have a standard structural-break model, the break fraction  $\tau = k_1/T$  can be consistently estimated. More interestingly, the estimated break has a convergence rate of  $T$ . That is,  $T(\hat{\tau} - \tau)$  is stochastically bounded despite measurement error in the regressors.

The above simple example highlights the main argument for consistency. We can extend this result to multiple-break models with general specification errors. Our estimation method does not assume that the number of breaks is known. It will be demonstrated that the estimated break point is consistent for one of the true break points when the model has multiple breaks, and that it has a convergence rate of  $T$  in spite of various misspecifications.

The paper is organized as follows: Section 2 presents the model and major assumptions. Section 3 shows the asymptotic behavior of the break-point estimator and the Sup-Wald statistic under misspecification. Simulations in support of our theories are given in Section

4. Section 5 provides an empirical application. The last section concludes the paper. All proofs are relegated to the appendix.

## 2. THE MODEL AND ASSUMPTIONS

To begin with, we present some frequently used mathematical notations in this paper.  $[x]$  denotes the greatest integer  $\leq x$ . The symbol ' $\xrightarrow{p}$ ' represents convergence in probability, ' $\xrightarrow{d}$ ' represents convergence in distribution, and ' $\implies$ ' signifies weak convergence in  $D[0, 1]$ , see Billingsley (1968) and Pollard (1984). All limits are defined as the sample size  $T \rightarrow \infty$  unless otherwise stated.

Consider a multivariate linear regression model with  $p$  break points at unknown time points, the number of break points  $p$  is also unknown. We assume that the true model has the following form:

$$y_t = f_1(x_{t1})\beta_{1i} + f_2(x_{t2})\beta_{2i} + \cdots + f_L(x_{tL})\beta_{Li} + u_t, \quad k_{i-1} < t \leq k_i \\ t = 1, 2, \dots, T; \quad i = 1, 2, \dots, p+1.$$

This is a model with  $p+1$  regimes, and the  $p$  break points are  $k_1, \dots, k_p$ . We define, throughout,  $k_0 = 0, k_{p+1} = T$ .

In matrix notation, the true model can be rewritten as

$$Y = \sum_{i=1}^{p+1} I_i F \beta_i + U, \quad (2.1)$$

where

$Y = (y_1 \ y_2 \ \dots \ y_T)'$  is a  $T$  by 1 vector of  $y_t$ ;

$F$  is a  $T$  by  $L$  matrix with  $(t, l)^{th}$  element  $f_l(x_{tl})$ , where  $f_l(\cdot)$  is a real-value function for  $l = 1, 2, \dots, L$ ;

$U = (u_1 \ u_2 \ \dots \ u_T)'$  is a  $T$  by 1 vector with the element  $u_t$  being a martingale difference sequence with  $\frac{1}{T} \sum_{t=1}^T u_t^2 \xrightarrow{p} \sigma^2 < \infty, \sup_t E |u_t|^{4+c} < \infty$  for some  $c > 0$  (Bai and Perron, 1998);

$\beta_i = (\beta_{1i} \ \beta_{2i} \ \dots \ \beta_{Li})'$  is a  $L$  by 1 vector of true parameters,  $i = 1, 2, \dots, p+1$ ;

$I_i$  is a  $T$  by  $T$  diagonal matrix with the  $(t, t)^{th}$  element being an indicator function  $1\{k_{i-1} < t \leq k_i\}$ ,  $i = 1, 2, \dots, p+1$ ,  $k_i$  are the dates of changes,  $i = 1, 2, \dots, p$ .

Let  $\tau_i = k_i/T$  be the true break fraction for  $i = 1, 2, \dots, p$ . Also let  $\tau_0 = 0, \tau_{p+1} = 1$ .

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

Without knowing the true model, the following regression model with a single break is estimated:

$$Y = I_a G \hat{\beta}_{a[\tau]} + I_b G \hat{\beta}_{b[\tau]} + \hat{V}, \quad (2.2)$$

where

$I_a$  is a  $T$  by  $T$  diagonal matrix with the  $(t, t)^{th}$  element being an indicator function  $1\{t < k\}$  with  $k = [T\tau]$ , and  $I_b = I_{T \times T} - I_a$ ;

$\hat{\beta}_{a[\tau]}$  and  $\hat{\beta}_{b[\tau]}$  are  $N$  by 1 vectors of OLS coefficient estimates for the first and second sample subsample respectively;

$G$  is a  $T$  by  $N$  matrix with the  $(t, n)^{th}$  element  $g_n(x_{tn})$ , where  $g_n(\cdot)$  is a real-value function,  $n = 1, 2, \dots, N$ ;

$\hat{V}$  is a  $T$  by 1 vector of the residuals for the misspecified model.

In this model, only one break point is estimated despite the existence of  $p$  true break points. Therefore, two kinds of misspecifications occur in this model: misspecification in regressors and misspecification in the number of breaks.

Define the following matrices:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E [f(x_t) f(x_t)'] = Q_{ff},$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E [g(x_t) g(x_t)'] = Q_{gg}$$

and

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E [g(x_t) f(x_t)'] = Q_{gf}.$$

where  $f(x_t) = (f_1(x_{t1}), \dots, f_L(x_{tL}))'$  is a  $L \times 1$  vector, and  $g(x_t) = (g_1(x_{t1}), \dots, g_N(x_{tN}))'$  is an  $N \times 1$  vector. We define  $Q_{fg} = Q_{gf}'$ .

$Q_{ff}$  and  $Q_{gg}$  are assumed to be of full rank. The matrix  $Q_{ff}$  being of full rank is a necessary identification condition even in the absence of breaks. The full rank requirement is standard. Define

$$Q = Q_{fg} Q_{gg}^{-1} Q_{gf}.$$

It is assumed that  $Q$  is of full rank<sup>2</sup>.

Following Chong (2003), we also assume, uniformly in  $\tau \in [0, 1]$ , that

$$(A1) \quad \frac{1}{T} F' I_a F \xrightarrow{p} \tau Q_{ff};$$

$$(A2) \quad \frac{1}{T} G' I_a G \xrightarrow{p} \tau Q_{gg};$$

$$(A3) \quad \frac{1}{T} G' I_a F \xrightarrow{p} \tau Q_{gf};$$

(A4)  $\det F' [I(j) - I(h)] F > 0$  and  $\det G' [I(j) - I(h)] G > 0$  for  $\forall j, h \in [0, T]$  and  $j > h$ , where  $I(j)$  and  $I(h)$  are  $T$  by  $T$  diagonal matrices with the  $(t, t)^{th}$  element being an indicator function  $1 \{t \leq j\}$  and  $1 \{t \leq h\}$  respectively;

(A5) There exist some real numbers  $r > 2$ ,  $C_l > 0$  and  $D_n > 0$  ( $1 \leq l \leq L, 1 \leq n \leq N$ ) such that for all  $0 \leq h < j < T$ , the two inequalities

$$E \left| \sum_{t=h+1}^j f_l(x_{tl}) u_t \right|^r \leq C_l (j-h)^{r/2}$$

and

$$E \left| \sum_{t=h+1}^j g_n(x_{tn}) u_t \right|^r \leq D_n (j-h)^{r/2}$$

hold;

$$(A6) \quad (\tau_1, \tau_2, \dots, \tau_p, \beta'_1, \beta'_2, \dots, \beta'_{p+1}) \in \Theta = [\underline{\tau}, \bar{\tau}]^p \times B^{p+1} \subset (0, 1)^p \times \mathbf{R}^{(p+1) \times L},$$

where  $B$  is a  $L$  dimension parameter space;

$$(A7) \quad \inf_i |\tau_{i+1} - \tau_i| > 0, \quad i = 0, 1, \dots, p, \quad \tau_0 = 0, \quad \tau_{p+1} = 1;$$

$$(A8) \quad \inf_i \|\beta_{i+1} - \beta_i\| > 0, \quad i = 0, 1, \dots, p.$$

<sup>2</sup>The assumption of full rank for  $Q$  is equivalent to  $Q_{gf}$  having a full column rank. We do not allow all the regressors to be completely unrelated to the true regressors. The requirement that  $Q_{gf}$  has full column rank says that at least a subset of regressors (same dimension as the true regressors  $f$ ) is correlated to the true regressors. This is a reasonable assumption since one cannot expect to consistently estimate the model with arbitrary unrelated regressors.

Assumptions (A1) - (A3) exclude trending regressors. Assumption (A4) guarantees the invertibility of the matrices defined there, and the OLS regression coefficients  $\widehat{\beta}_a$  and  $\widehat{\beta}_b$  can thus be estimated. The uniform convergence result is ensured by assumption (A5). Assumption (A6) states that the true break points are in a compact set in  $(0, 1)$  as the estimated  $\widehat{\beta}_a$  and  $\widehat{\beta}_b$  are not defined at the boundary of the time domain. Assumption (A7) requires two consecutive change points to be separated far enough from each other. Assumption (A8) states that the magnitudes of changes should not be negligible.

In model (2), there are two kinds of misspecifications, namely,  $F \neq G$  or the number of true breaks is more than one ( $p > 1$ ). In other words, the model is misspecified if the functional form of the regressors or the true number of changes is actually more than one.

### 3. ASYMPTOTICS

#### 3.1. Consistency

Following Chong (2001), we define

$$\widehat{\tau}_T = \arg \min_{\tau \in [\underline{\tau}, \overline{\tau}]} RSS(\tau, 0, 1), \quad (3.3)$$

where  $RSS(\tau, \tau^v, \tau^l)$  denotes the residual sum of squares of model (2) regressed on data dating from  $\tau^v T$  to  $\tau^l T$ , and

$$RSS(\tau, 0, 1) = \left\| Y - I_a G \widehat{\beta}_{a[\tau]} - I_b G \widehat{\beta}_{b[\tau]} \right\|^2$$

is the residual sum of squares for the whole sample. For any given value of  $\tau$ , the least squares estimators of the pre- and post-shift parameters of model (2) are

$$\begin{aligned} \widehat{\beta}_{a[\tau]} &= (G' I_a G)^{-1} G' I_a Y, \\ \widehat{\beta}_{b[\tau]} &= (G' I_b G)^{-1} G' I_b Y. \end{aligned}$$

For  $\tau \in [\tau_h, \tau_{h+1}]$ ,  $h = 0, 1, \dots, p$ ,

$$\begin{aligned} \widehat{\beta}_{a[\tau]} &\xrightarrow{p} Q_{gg}^{-1} Q_{gf} \Gamma_1(\tau) \\ \widehat{\beta}_{b[\tau]} &\xrightarrow{p} Q_{gg}^{-1} Q_{gf} \Gamma_2(\tau) \end{aligned} \quad (3.4)$$

where for  $\tau \in [\tau_h, \tau_{h+1}]$ ,

$$\begin{aligned} \Gamma_1(\tau) &= \beta_1 \quad \text{for } h = 0 \\ &= \frac{\sum_{i=1}^h (\tau_i - \tau_{i-1}) \beta_i + (\tau - \tau_h) \beta_{h+1}}{\tau}, \quad \text{for } h = 1, 2, \dots, p \end{aligned} \quad (3.5)$$

$$\begin{aligned} \Gamma_2(\tau) &= \frac{\sum_{i=h+2}^{p+1} (\tau_i - \tau_{i-1}) \beta_i + (\tau_{h+1} - \tau) \beta_{h+1}}{1 - \tau}, \quad \text{for } h = 0, 1, \dots, p-1 \\ &= \beta_{p+1}, \quad \text{for } h = p \end{aligned} \quad (3.6)$$

In general, for any interval  $(\tau^v, \tau^l) \subseteq [0, 1]$ , we have

$$\frac{1}{T} RSS(\tau, \tau^v, \tau^l) \xrightarrow{p} R(\tau, \tau^v, \tau^l) \quad \text{uniformly,}$$

where  $R(\tau, \tau^v, \tau^l)$  is a piecewise concave function of  $\tau$  defined on  $(\tau^v, \tau^l)$ .

In particular, it can be shown that

$$\frac{1}{T}RSS(\tau, 0, 1) \xrightarrow{p} R(\tau, 0, 1), \tau \in [0, 1]$$

uniformly, where

$$R(\tau, 0, 1) = \sigma^2 + \sum_{i=1}^{p+1} (\tau_i - \tau_{i-1}) \beta'_i Q_{ff} \beta_i - \tau \Gamma'_1(\tau) Q \Gamma_1(\tau) - (1 - \tau) \Gamma'_2(\tau) Q \Gamma_2(\tau).$$

Note that for  $\tau \in [0, \tau_1)$ , we have

$$\frac{\partial R(\tau, 0, 1)}{\partial \tau} = - \left( \frac{1}{1 - \tau} \right)^2 L'_1 Q L_1 \leq 0$$

and

$$\frac{\partial^2 R(\tau, 0, 1)}{\partial \tau^2} = - \frac{2}{(1 - \tau)^3} L'_1 Q L_1 \leq 0,$$

where

$$L_1 = \sum_{i=2}^{p+1} (\tau_i - \tau_{i-1}) \beta_i - (1 - \tau_1) \beta_1.$$

For  $\tau \in [\tau_h, \tau_{h+1})$ ,  $h = 1, 2, \dots, p - 1$ , we have

$$\frac{\partial R(\tau, 0, 1)}{\partial \tau} = \Gamma'_1(\tau) Q \Gamma_1(\tau) - \Gamma'_2(\tau) Q \Gamma_2(\tau) - 2\beta'_{h+1} Q (\Gamma_1(\tau) - \Gamma_2(\tau)),$$

$$\begin{aligned} \frac{\partial^2 R(\tau, 0, 1)}{\partial \tau^2} &= -2 \frac{1}{\tau} (\beta_{h+1} - \Gamma_1(\tau))' Q (\beta_{h+1} - \Gamma_1(\tau)) \\ &\quad - 2 \frac{1}{1 - \tau} (\beta_{h+1} - \Gamma_2(\tau))' Q (\beta_{h+1} - \Gamma_2(\tau)) \leq 0. \end{aligned}$$

For  $\tau \in [\tau_p, 1]$ , we have

$$\frac{\partial R(\tau, 0, 1)}{\partial \tau} = \frac{1}{\tau^2} L'_2 Q L_2 \geq 0,$$

where

$$L_2 = \sum_{i=1}^p (\tau_i - \tau_{i-1}) \beta_i - \tau_p \beta_{p+1},$$

and

$$\frac{\partial^2 R(\tau, 0, 1)}{\partial \tau^2} = - \frac{2}{\tau^3} L'_2 Q L_2 \leq 0.$$

The beauty of  $R(\tau, 0, 1)$  is that it is concave within any interval  $[\tau_h, \tau_{h+1}]$ ,  $h = 0, 1, \dots, p$ . In addition, for the two boundary intervals, we have  $\frac{\partial R(\tau, 0, \tau_1)}{\partial \tau} \leq 0$  and  $\frac{\partial R(\tau, \tau_p, 1)}{\partial \tau} \geq 0$ , implying that the function is decreasing for  $\tau \in [0, \tau_1]$  and is increasing for  $\tau \in [\tau_p, 1]$ . Thus, the function  $R(\tau, 0, 1)$  achieves its minimum at one of the true break points  $(\tau_1, \dots, \tau_p)$ , which in turn implies that  $\hat{\tau}_T$  is consistent for one of the true break points.

**PROPOSITIONS 3.1.** *Under assumptions (A1)-(A8),  $R(\tau, \tau^v, \tau^l)$  defined in any subsample  $[\tau^v, \tau^l] \subseteq [0, 1]$  is piece-wise concave.*

The proof is provided in the appendix. Proposition 1 suggests that the estimated

break point is consistent for one of the true breaks. The piece-wise concavity over each subinterval guarantees the consistent estimation of all the break points. The following theorem strengthens the consistency argument by including the convergence rate for the estimated break points.

**THEOREM 3.1.** *Under assumptions (A1) – (A8), for some  $k=1,2,\dots,p$*

(i)  $\widehat{\tau}_T$  is consistent for one of the true break points, i.e.,

$$\text{plim } \widehat{\tau}_T = \tau_k.$$

(ii) Suppose that  $\widehat{\tau}_T$  is consistent for  $\tau_k$ . Then

$$T(\widehat{\tau}_T - \tau_k) = O_p(1).$$

**Proof of Theorem 3.1:** Recall the true model has the form:

$$Y = \sum_{i=1}^{p+1} I_i F \beta_i + U. \quad (3.7)$$

Let  $F_t = f(x_t) = (f_1(x_{t1}), \dots, f_L(x_{tL}))'$  and similarly, let  $G_t = g(x_t)$ . Thus

$$F = (F_1, F_2, \dots, F_T)',$$

$$G = (G_1, G_2, \dots, G_T)'.$$

Project  $F_t$  on  $G_t$ , we have

$$F_t = CG_t + e_t, \quad (3.8)$$

where  $C$  is a  $L \times N$  matrix of coefficient. Note that the rank of  $C$  is also  $L$ , which is less than or equal to  $N$ . This follows from  $\frac{1}{T} \sum_{t=1}^T F_t G_t' = C \frac{1}{T} \sum_{t=1}^T G_t G_t' + \frac{1}{T} \sum_{t=1}^T e_t G_t'$ . From the orthogonality of  $G_t$  and  $e_t$  due to the projection, we have, in the limit,  $Q_{fg} = CQ_{gg}$ . Because the rank of  $Q_{fg}$  is  $L$  and the rank of  $Q_{gg}$  is  $N$  ( $N \geq L$ ), the rank of  $C$  must be  $L$  (full row rank).

Let  $e = (e_1, e_2, \dots, e_T)'$ . The model can be rewritten as

$$Y = \sum_{i=1}^{p+1} I_i (GC' + e) \beta_i + U = \sum_{i=1}^{p+1} I_i G \theta_i + V, \quad (3.9)$$

where

$$\theta_i = C' \beta_i$$

and

$$V = \sum_{i=1}^{p+1} I_i e \beta_i + U.$$

Furthermore, from  $\beta_{i+1} - \beta_i \neq 0$ , we have  $\theta_{i+1} - \theta_i = C'(\beta_{i+1} - \beta_i) \neq 0$  because  $C' = Q_{gg}^{-1} Q_{gf}$  has a rank of  $L$  and  $\beta_i$  is  $L \times 1$ . Thus, the model reduces to that of a standard multiple-change model with  $G$  being considered as the true regressors and  $\theta_i$  as the new coefficients. From this point of view, the estimated model only has misspecification in the number of breaks points. Now invoke the results of Bai (1997),  $\widehat{\tau}_T$  is T-consistent for one of the true breaks. This completes the proof of Theorem 3.1  $\square$ .<sup>3</sup>

<sup>3</sup>The proof uses the fact that  $Q_{gf}(\beta_{i+1} - \beta_i) \neq 0$ , which is a necessary condition for consistent estimation of the break point. This is because the condition implies the existence of a break point in the

Theorem 3.1 shows the consistency of the estimated change point for one of the true ones. However, which of the change points will be identified depends on a number of factors. It is not necessarily the case that the break with the biggest magnitude will be identified first. The result also depends on the duration of the break and the form of misspecification. For example, in the simplest case of Chong (1995), where  $F = G$  and the true model has two breaks, but a single-break model is estimated, if

$$\frac{(\beta_3 - \beta_2)' Q (\beta_3 - \beta_2)}{(\beta_2 - \beta_1)' Q (\beta_2 - \beta_1)} < \frac{\tau_1 (1 - \tau_1)}{\tau_2 (1 - \tau_2)},$$

then  $\hat{\tau} \xrightarrow{p} \tau_1$ , and if this inequality is reversed, then  $\hat{\tau} \xrightarrow{p} \tau_2$ .

According to Theorem 3.1, we can first consistently estimate one of the break points  $\hat{\tau}_T$  in the whole sample, then split the sample into pre- and post-shift subsample at the break fraction  $\hat{\tau}_T$ . Within both subsamples  $[0, \hat{\tau}_T]$  and  $(\hat{\tau}_T, 1]$ , estimate  $\hat{\tau}'_T = \arg \min RSS(\tau, 0, \hat{\tau}_T)$  and  $\hat{\tau}''_T = \arg \min RSS(\tau, \hat{\tau}_T, 1)$  to obtain the next two break points, assuming each subinterval contains a break point. This process continues until all other break points are located<sup>4</sup>. According to Proposition 1, the estimated break is consistent. Alternatively, in each subsequent round of estimation, only one break point is chosen in one of the subsamples for which the reduction in residual sum of squares is greatest. This method is considered by Bai and Perron (1998) and is referred to as the sequential method. The process stops when the specified number of breaks is obtained.

### 3.2. The number of break points and the Wald test

Consistent estimation of a break point does not hinge on the prior knowledge of the number of breaks, as long as there exists at least one break point. However, if the objective is to estimate all break points, clearly, one needs to know the number of break points. Recently, Altissimo and Corradi (2003) propose a consistent estimation method for the number of break points when the model is correctly specified. In this paper, we discuss three estimation methods. The first is based on the Sup-Wald test procedure of Chong (2003). The second estimation method is based on a Wald test with a heteroskedastic and autocorrelation (HAC) robust estimator of the variance. The third estimation is to use the information criteria of Yao (1998).

**3.2.1. The Sup-Wald test statistic** The Sup-Wald test statistic together with sample-splitting method can be used to determine the location and number of the breaks. Suppose model (2.1) is the true model, but model (2) is estimated; the Sup-Wald statistic for the null hypothesis  $H_0 : \beta_1 = \beta_2 = \dots = \beta_{p+1} = \beta$  is defined as:

projected model. The condition that  $Q_{gf}(\beta_{i+1} - \beta_i) \neq 0$  ( $i = 1, 2, \dots, m$ ) does not always require  $Q_{gf}$  to be of full column rank for a given parameter configuration  $\beta$ . However, in order for the condition to hold for all configurations of  $\beta$ , the matrix  $Q_{gf}$  must be of full column rank. Because Theorem 1 makes a general statement that covers all configurations of  $\beta$ , the full column rank requirement for  $Q_{gf}$  becomes necessary.

<sup>4</sup>Note that  $R(\tau, 0, 1)$  will no longer be piecewise concave and the consistency of the estimation is not guaranteed if  $Q_{fg} = 0$ . For example, when  $f(x_t) = x_t \sim i.i.d.N(0, \sigma_x^2)$  and  $g(x_t) = x_t^2$ , we have  $Q_{fg} = 0$ , thus  $RSS(\tau, 0, 1) \xrightarrow{p} R(\tau, 0, 1) = \sigma^2 + \sum_{i=1}^{p+1} (\tau_i - \tau_{i-1}) \beta_i' Q_{ff} \beta_i$  which is no longer a function of  $\tau$ .



$$\sup_{\tau \in \mathbf{S}} W_T(\tau, 0, 1) = \sup_{\tau \in \mathbf{S}} \frac{T\tau(1-\tau)}{RSS(\tau, 0, 1)} \left( \hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]} \right)' G' G \left( \hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]} \right), \quad (3.10)$$

where  $\mathbf{S}$  is a set with closure in  $[0, 1]$ .

The second and the third parameter of  $W_T(\tau, 0, 1)$  imply that the test is calculated with the full sample data. As shown in Chong (2003), although the pre- and post-shift estimators will not be consistent in the presence of misspecification, they will converge to the same constant under the null hypothesis of no structural break. Meanwhile, if there are structural breaks, the probability limits of the pre- and post-shift estimators will generally be different in the presence of specification errors. Therefore, the Wald type test based on the magnitude of the estimated break will still be a consistent test.

To derive the distribution of  $W_T(\tau, 0, 1)$ , we make the following assumptions:

- (B1)  $u_t$  follows an *i.i.d*  $(0, \sigma^2)$  with  $\sigma^2 < \infty$ ;
- (B2)  $\frac{1}{\sqrt{T}} F' I_a U \implies B_{fu}(\tau)$ , where  $B_{fu}(\tau)$  is a  $L$ -vector zero-mean Gaussian process with variance  $\tau\sigma^2 Q_{ff}$ ;
- (B3)  $\frac{1}{\sqrt{T}} G' I_a U \implies B_{gu}(\tau)$ , where  $B_{gu}(\tau)$  is a  $N$ -vector zero-mean Gaussian process with variance  $\tau\sigma^2 Q_{gg}$ ;
- (B4)  $\sqrt{T} \text{Vec} \left( \frac{1}{T} G' I_a G - \tau Q_{gg} \right) \implies B_{gg}(\tau)$ , where  $B_{gg}(\tau)$  is a  $N^2 \times 1$  vector of zero-mean Gaussian processes with covariance matrix

$$\tau\sigma^2 \lim_{T \rightarrow \infty} TE \left[ \text{Vec} \left( \frac{1}{T} G' G - Q_{gg} \right) \left( \text{Vec} \left( \frac{1}{T} G' G - Q_{gg} \right) \right)' \right];$$

- (B5)  $\sqrt{T} \text{Vec} \left( \frac{1}{T} G' I_a F - \tau Q_{fg} \right) \implies \text{Vec} B_{fg}(\tau)$ , where  $B_{fg}(\tau)$  is a  $NL \times 1$  matrix of zero-mean Gaussian processes with variance

$$\tau\sigma^2 \lim_{T \rightarrow \infty} TE \left[ \text{Vec} \left( \frac{1}{T} G' F - Q_{fg} \right) \left( \text{Vec} \left( \frac{1}{T} G' F - Q_{fg} \right) \right)' \right].$$

Chong (2003) derives the null distribution of  $W_T(\tau, 0, 1)$  under assumptions (A1) - (A5) and (B1) - (B5).

**THEOREM 3.2.** *Under assumptions (A1) - (A5) and (B1) - (B5), and  $H_0 : \beta_1 = \beta_2 = \dots = \beta_{p+1} = \beta$ , as  $T \rightarrow \infty$ , we have*

$$W_T(\tau, 0, 1) \implies \frac{[\tau B_A(1) - B_A(\tau)]' Q_{gg}^{-1} [\tau B_A(1) - B_A(\tau)]}{\tau(1-\tau) [\sigma^2 + \beta' (Q_{ff} - Q) \beta]}$$

and

$$\sup_{\tau \in \mathbf{S}} W_T(\tau, 0, 1) \xrightarrow{d} \sup_{\tau \in \mathbf{S}} \frac{[\tau B_A(1) - B_A(\tau)]' Q_{gg}^{-1} [\tau B_A(1) - B_A(\tau)]}{\tau(1-\tau) [\sigma^2 + \beta' (Q_{ff} - Q) \beta]}$$

where  $\mathbf{S}$  is a set with closure in  $[\underline{\tau}, \bar{\tau}]$ ,  $B_A(\tau) = [B_{fg}(\tau) - B_{gg}(\tau) Q_{gg}^{-1} Q_{fg}] \beta + B_{gu}(u)$  is a  $N$ -vector of Brownian motions on  $[0, 1]$  and  $Q = Q'_{fg} Q_{gg}^{-1} Q_{fg}$ .

The proof is given in Chong (2003) and is thus omitted. If there is no specification error,

we have  $B_A(\tau) = B_{fu}(u)$ ,  $Q = Q_{ff}$ , and  $\sup_{\tau \in \mathbf{S}} W_T(\tau, 0, 1) \xrightarrow{d} \sup_{\tau \in \mathbf{S}} \frac{\|\tau B(1) - B(\tau)\|^2}{\tau(1-\tau)}$  under  $H_0$ , where  $B(\tau)$  is a  $L$ -vector of independent Brownian motions. With little modification, Theorem 2 can also be shown to hold in any subsample. To estimate the number of breaks, one can follow the sample splitting method of Chong (2001). After the first break point is identified, we split the sample into two subsamples using the first break-point estimate as a cut-off point. Then the same test is performed on the two subsamples. The procedure continues until all the break points are identified.

Note that the distribution of Sup-Wald statistic is now affected by the form of misspecification in an unknown manner. However, if the misspecification is due to measurement errors, we can still perform the test in some special cases. To see this, let the true model be:

$$Y = \sum_{i=1}^{p+1} I_i F \beta_i + U,$$

but we estimate

$$Y = I_a(F + \Delta) \hat{\beta}_{a[\tau]} + I_b(F + \Delta) \hat{\beta}_{b[\tau]} + \hat{V},$$

where  $\Delta = (\delta_1, \delta_2, \dots, \delta_T)'$  is a  $T$  by  $L$  matrix and  $\delta_t = (\delta_{t1}, \delta_{t2}, \dots, \delta_{tL})'$ ,  $(t = 1, 2, \dots, T)$ , is the measurement error vector. Assuming that  $\{\delta_{tl}\}_{t=1}^T$  are zero-mean *i.i.d* across  $t$  and  $l$ , and are independent of  $F$  and  $U$ , then we have

$$E(\Delta) = 0$$

and

$$\lim_{T \rightarrow \infty} \frac{1}{T} \Delta' \Delta = \Sigma,$$

where  $\Sigma$  is a  $L \times L$  diagonal matrix with the  $(l, l)^{th}$  element  $var(\delta_{tl}) = \sigma_l^2$ . Note that in this case,  $Q_{fg} = Q_{ff}$  and  $Q_{gg} = Q_{ff} + \Sigma$ . It is not clear how  $\Delta$  affects the test statistic when  $Q_{ff}$  is not diagonal. However, when  $L = 1$  and if  $f(x_t) \sim \text{i.i.d.} N(0, \alpha_f \sigma^2)$ ,  $\delta_t \sim \text{i.i.d.} N(0, \alpha_\delta \sigma^2)$ , where  $\alpha_f$  and  $\alpha_\delta$  are the ratios of  $var[f(x_t)]$  and  $var(\delta_t)$  to that of  $u_t$ , then  $\sup_{\tau \in \mathbf{S}} W_T(\tau, 0, 1) \xrightarrow{d} \sup_{\tau \in \mathbf{S}} \frac{(\tau B(1) - B(\tau))^2}{\tau(1-\tau)}$ , which is the conventional null distribution of Sup-Wald test statistic for testing structural break.

*3.2.2. The HAC adjusted Wald test statistic* In general, the limiting distribution of the Sup-Wald statistic depends on unknown parameters, except for some special cases. The underlying reason is that misspecification and measurement error induce heteroskedasticity in the sense that  $E(g_t g_t' v_t^2) \neq E(g_t g_t') E v_t^2$ . As is well known, the usual F test (and hence the Sup-Wald test) fails under heteroskedasticity. Thus, a heteroskedasticity-robust covariance should be used. In this paper, we propose a modified test whose asymptotic distribution will not be affected by the functional form  $F$  and the unknown parameter  $\beta$ . As the misspecification in the functional form can be absorbed by the error term, we can construct a HAC adjusted Wald test statistic (Bai and Perron, 1998) as follows:

$$W_T^{HAC}(\tau, 0, 1) = T\tau(1-\tau) \left( \hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]} \right)' \hat{\Omega}^{-1} \left( \hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]} \right), \quad (3.11)$$

where  $S$  is a set with closure in  $[0, 1]$ , and the estimated asymptotic variance of  $\hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]}$  is  $\frac{1}{\tau(1-\tau)}\hat{\Omega}$ , where

$$\hat{\Omega} = p \lim \left( \frac{G'G}{T} \right)^{-1} \left[ \frac{\sum_{t=1}^T g_t g_t' \hat{v}_t^2}{T} \right] \left( \frac{G'G}{T} \right)^{-1},$$

and  $\hat{v}_t = y_t - g_t' \hat{\beta}$  ( $t = 1, 2, \dots, T$ ) are the estimated residuals. Note that  $\hat{\Omega}$  is robust to heteroskedasticity. One can also employ the Newey-West (1987) estimator to make it robust to serial correlation as well.

**THEOREM 3.3.** *Under assumptions (A1) - (A5) and (B1) - (B5), and  $H_0 : \beta_1 = \beta_2 = \dots = \beta_{p+1} = \beta$ , as  $T \rightarrow \infty$ , we have  $W_T^{HAC}(\tau, 0, 1) \xrightarrow{d} \chi_N^2$  for each fixed  $\tau$ , and  $\sup_{\tau \in S} W_T^{HAC}(\tau, 0, 1) \xrightarrow{d} \sup_{\tau \in S} \frac{\|\tau B(1) - B(\tau)\|^2}{\tau(1-\tau)}$ , where  $S$  is a set with closure in  $[\underline{\tau}, \bar{\tau}]$ ,  $B(\tau)$  is a  $N$ -vector of independent Brownian motions.*

**Proof of Theorem 3.3** See the Appendix.

According to Theorem 3.3,  $\sup W_T^{HAC}(\tau, 0, 1)$  converges to the same asymptotic distribution of the Sup-Wald test of Andrews (1993).

*3.2.3. Information criteria of Yao (1988)* Alternatively, one can use the Schwarz information criterion approach of Yao (1988) to determine the total number of breaks. Once the number of breaks is determined, the dynamic programming method of Bai and Perron (1998) can be used to estimate all breaks simultaneously. No matter which method is used, the estimated break point is  $T$  consistent.

## 4. SIMULATIONS

**Experiment 1.** This experiment shows the distribution of the first break-point estimator in a two-break model under misspecifications. We perform the following simulations:

True model:

$$\begin{aligned} y_t &= 1 + x_t + u_t & \text{for } t \in [1, \frac{T}{3}] \\ y_t &= 1 + 3x_t + u_t & \text{for } t \in (\frac{T}{3}, \frac{2T}{3}] \\ y_t &= 1 + 4x_t + u_t & \text{for } t \in (\frac{2T}{3}, T] \end{aligned}$$

Estimated model:

$$\begin{aligned} y_t &= \hat{\beta}_{01} + \hat{\beta}_1 g(x_t) + v_t & \text{for } t \leq \tau T \\ y_t &= \hat{\beta}_{02} + \hat{\beta}_2 g(x_t) + v_t & \text{for } t > \tau T \end{aligned}$$

Model	(1)	(2)	(3)
$g(x_t)$	$x_t$	$x_t^3$	$\ln x_t$

where  $T = 1000$ ;  $u_t \sim i.i.d.N(0, 1)$ ;  $x_t \sim i.i.d.U(1, 10)$ ;  $\{x_t\}_{t=1}^T$  and  $\{u_t\}_{t=1}^T$  are independent of each other. Table 1 shows the frequency distribution of the minimum of  $RSS(\tau, 0, 1)$  around the two true break points. The number of replication is set to

**Table 1.** Distribution of the first estimated break point

$\hat{k}_1 - 334$	Estimated Models		
	(1)	(2)	(3)
< -6	0	10	0
-6	0	3	0
-5	0	14	0
-4	0	17	0
-3	2	29	0
-2	7	68	0
-1	81	126	2
0	820	429	446
1	66	112	189
2	17	56	111
3	3	47	72
4	4	26	37
5	0	34	34
6	0	11	26
> 6	0	18	83

$N = 1000$ . Note that the first break will be identified first. Table 1 shows that the first break points always have a high chance of being identified.

**Experiment 2.** When there is only one regressor in both the true and the estimated models, we have shown that the Sup-Wald statistic will not be affected asymptotically. This experiment studies the behavior of Sup-Wald statistic in the presence of measurement error. We estimate the following model:

True model:

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

Misspecified model:

$$y_t = \hat{\beta}(x_t + \delta_t) + v_t, \quad t = 1, 2, \dots, T.$$

$T = 1000$  (sample size);

$N = 10000$  (number of replications)

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

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

$\delta_t \sim i.i.d.N(0, \alpha_\delta)$ ;

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

Using the misspecified model, we simulate the critical values of null hypothesis of no break point under measurement error. Table 2 reports the critical value  $c$  such that  $\Pr \left[ \sup_{\tau \in (\lambda, 1-\lambda)} W_T(\tau, 0, 1) > c \right] = \alpha$ . The values are compared with Andrews (1993).

It can be clearly observed from Table 2 that the critical values are very close to those of Andrews (1993). Additional simulations (not reported here) show that, under general

**Table 2.** Critical value Table:  $\beta = 1$ 

$\lambda \backslash \alpha$	Andrews(1993)			$\alpha_\delta = 0$			$\alpha_\delta = 1$			$\alpha_\delta = 2$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
0.1	7.63	9.31	12.69	7.60	9.51	12.55	7.55	9.19	13.06	7.51	8.21	12.80
0.2	6.80	8.45	11.69	6.74	8.41	11.79	6.74	8.42	11.58	6.69	8.21	11.62
0.3	6.05	7.51	10.91	6.15	7.44	10.80	5.87	7.46	11.08	6.13	7.46	11.19
0.4	5.10	6.57	9.82	5.18	6.50	9.77	4.90	6.35	9.81	4.94	6.34	10.11

**Table 3.** Comparison results for different models and methods

Model	SEQ with BIC	GLM with BIC	SEQ with Sup-HAC-Wald
<b>Mean-shift</b>	1st break: 79	2 breaks: 47, 79	1st break: 79
	2nd break: 47		2nd break: 47
<b>AR(1)</b>	1st break: 81	2 breaks: 46, 78.	1st break: 81; 2nd
	2nd break: 46		break: 46; 3rd break: 24
<b>AR(2)</b>	1st break: 80	1 break: 80	1st break: 80; 2nd
			break: 46; 3rd break: 24

misspecifications, with HAC-adjusted Sup-Wald test, the critical values are also close to those of Andrews (1993), as predicted by Theorem 3.

## 5. EMPIRICAL APPLICATION

In this section, we provide an application on the U.S. ex-post real interest rate series considered by Garcia and Perron (1996) and Bai and Perron (2003). The real interest rate series is constructed from the three-month treasury bill rate deflated by the CPI taken from the Citibase data base. Quarterly data from 1961Q1 to 1986Q3 are used. Bai and Perron (2003) apply a simple mean-shift model to the data and found two breaks. In this paper, we check the robustness of the estimation when different models are used. Meanwhile, we would also like to see whether the sequential estimation method (SEQ) will give the same estimation result as the global minimization method (GLM). The latter estimates all the break points simultaneously. Two methods are used to determine the number of breaks. The first is the BIC of Yao (1988), while the second one is the Sup-HAC weighted Wald test.

Note that the sequential estimation method can obtain similar results as the global minimization method for all cases. For the mean-shift model, they give identical results. Using the BIC, two breaks are detected for the mean-shift model and for the AR(1) model, and only one break has been found for the AR(2) model. However, if we use the Sup-HAC-Wald test, we obtain two breaks for the mean-shift model and three for the AR(1) and the AR(2) models. The number of breaks detected by the BIC and Sup-HAC-Wald test can vary across different models, which is not unexpected in a finite sample. According to the Sup-HAC weighted Wald test, we conclude that there are three breaks in the AR model. Note that the estimation results are very close regardless of the method used.

## 6. CONCLUSION

In this paper, we combine the works of Bai (1997) and Chong (2003) to examine the consistency of the break-point estimator under misspecification of the independent variables in a multiple-break model with unknown number of break points. The misspecification can occur in different dimensions: (i) the number of breaks is incorrectly specified, (ii) the functional forms of regressors are incorrectly specified, (iii) variables are observed with errors. It is shown that the residual sum of squares divided by the sample size converges uniformly to a piecewise concave function, which in turn implies that, if a one-break model is estimated in the presence of multiple breaks, the break-point estimator will converge to one of the true break points. Thus, it is possible to locate other break points by using the sample splitting method. The rate of convergence for the estimated break points is also provided. If the number of break points is known, all break points can be consistently estimated one by one. The issue of how to determine the number of breaks is also discussed. Experimental and empirical evidence confirms our results.

## ACKNOWLEDGEMENTS

The authors would like to thank Stéphane Grégoir, Win lin Chou, Andy Kwan, Henry Lin and two anonymous referees for helpful comments. All errors are ours.

## REFERENCES

- Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821-856.
- Altissimo, F. and V. Corradi (2003). Strong rules for detecting the number of breaks in a time series. *Journal of Econometrics* 117, 207-244.
- Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory* 13, 315-352.
- Bai, J. and P. Perron (1998). Testing for and estimating of multiple structural changes. *Econometrica* 66, 47-79.
- Bai, J. and P. Perron (2003). Computation and analysis of multiple structural-change models. *Journal of Applied Econometrics* 18, 1-22.
- Billingsley, P. (1968). *Convergence of Probability Measures*. New York: John Wiley and Sons.
- Busetti, F. and A. Harvey (1998). Testing for the presence of a random walk in series with structural breaks. *Discussion Paper No. Em/98/365*, London School of Economics and Political Science.
- Chang, Y.P. and W.T. Huang (1997). Inference for the linear errors-in-variables with change point models. *Journal of the American Statistical Association* 92, 171-178.
- Chesher, A. (1991). The effect of measurement error. *Biometrika* 78, 451-462.
- Chong, T.T.L. (1995). Partial parameter consistency in a misspecified structural change model. *Economics Letters* 49, 351-357.
- Chong, T.T.L. (2001). Estimating the locations and number of change points by the sample-splitting method. *Statistical Papers* 42, 53-79.
- Chong, T.T.L. (2003). Generic consistency of the break-point estimator under specification errors. *Econometrics Journal* 6, 167-192.

- Forchini, G. (2002), Optimal similar tests for structural change for the linear regression model. *Econometric Theory* 18, 853-867.
- Garcia, R. and P. Perron (1996), An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics* 78, 111-125.
- Lee, J., Huang, C.J and Y.C. Shin (1997). On stationary tests in the presence of structural breaks. *Economics Letters* 55, 165-172.
- Nelson, D.B. (1995). Vector attenuation bias in the classical errors-in-variables model. *Economics Letters* 49, 345-349.
- Newey, W. K. and K.D. West (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.
- Perron, P. (2006). Dealing with structural breaks, in *Palgrave Handbook of Econometrics*, Vol. 1: Econometric Theory, K. Patterson and T.C. Mills (eds.), Palgrave Macmillan, 278-352.
- Pollard, D. (1984). *Convergence of Stochastic Process*. New York: Springer Inc.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica* 48, 817-838.
- Yao, Y.C. (1988). Estimating the number of change-points via Schwarz criterion. *Statistics & Probability Letters* 6, 181-189.

## APPENDIX

**Proof of Proposition 1:** We shall establish the piece-wise concavity of  $R(\tau, \tau^v, \tau^l)$ .

Suppose there are  $q$  ( $q \leq p$ ) breaks within the segment  $t = v + 1, \dots, l - 1, l$ . Let:

- (i)  $\tau^v = v/T, \tau^l = l/T, [\tau^v, \tau^l] \subset [\underline{\tau}, \bar{\tau}]$ ;
- (ii)  $\tau^v < \bar{\tau}_1 < \bar{\tau}_2 < \dots < \bar{\tau}_{q-1} < \bar{\tau}_q < \tau^l$ , where  $\bar{\tau}_i, i = 1, 2, \dots, q$  are  $q$  consecutive true change points with  $\bar{\tau}_1 = \tau_j, \forall j \in \{1, 2, \dots, p - q\}$ ;
- (iii) Let  $I^v, I^l, I_a, I_b, I_i$  be respectively a  $(l - v) \times (l - v)$  diagonal matrix with the  $(t - v, t - v)^{th}$  element an indicator function  $1\{v < t \leq \bar{\tau}_1 T\}, 1\{\bar{\tau}_q T < t \leq l\}, 1\{v < t \leq \tau T\}, 1\{\tau T < t \leq l\}, 1\{\bar{\tau}_{i-1} T < t \leq \bar{\tau}_i T\}, i = 2, \dots, q$ .

For convenience in notation, let  $\bar{\tau}_0 = \tau^v, \bar{\tau}_{q+1} = \tau^l, I_1 = I^v$ , and  $I_{q+1} = I^l$ .

The true model in this segment is

$$Y = \sum_{i=1}^{q+1} I_i F \beta_{r-1+i} + U. \quad (\text{A.1})$$

where  $r$  is an integer such that  $\tau_{r-1} \leq \tau_v < \tau_r$ .

Let  $\bar{\beta}_i = \beta_{r-1+i}, i = 1, 2, \dots, q + 1$ . We then rewrite (12) as:

$$Y = \sum_{i=1}^{q+1} I_i F \bar{\beta}_i + U.$$

Due to misspecification, model (2) is estimated.

For  $\bar{\tau}_0 \leq \tau < \bar{\tau}_1$ , we have

$$I_a I_1 = I_a;$$

$$I_a I_i = 0, i = 2, 3, \dots, q + 1;$$

$$I_b I_1 = (I - I_a) I_1 = I_1 - I_a;$$

$$I_b I_i = I_i, i = 2, 3, \dots, p.$$

Thus,

$$\begin{aligned}\hat{\beta}_{a[\tau]} &= (G' I_a G)^{-1} G' I_a Y \\ &= \left( \frac{G' I_a G}{T'} \right)^{-1} \frac{1}{T'} G' I_a \left( \sum_{i=1}^{q+1} I_i F \bar{\beta}_i + U \right) \\ &\xrightarrow{p} [(\tau - \bar{\tau}_0) Q_{gg}]^{-1} (\tau - \bar{\tau}_0) Q_{gf} \bar{\beta}_1 \\ &= Q_{gg}^{-1} Q_{gf} \bar{\beta}_1,\end{aligned}$$

$$\begin{aligned}\hat{\beta}_{b[\tau]} &= (G' I_b G)^{-1} G' I_b Y = \left( \frac{G' I_b G}{T'} \right)^{-1} \frac{1}{T'} G' I_b \left( \sum_{i=1}^{q+1} I_i F \bar{\beta}_i + U \right) \\ &\xrightarrow{p} (\bar{\tau}_{q+1} - \tau) Q_{gg}^{-1} Q_{gf} \left[ (\bar{\tau}_1 - \tau) \bar{\beta}_1 + \sum_{i=1}^q (\bar{\tau}_{i+1} - \bar{\tau}_i) \bar{\beta}_{i+1} \right] \\ &= Q_{gg}^{-1} Q_{gf} \Gamma_2(\tau),\end{aligned}$$

uniformly, where

$$\Gamma_2(\tau) = \frac{1}{\bar{\tau}_{q+1} - \tau} \left[ (\bar{\tau}_1 - \tau) \bar{\beta}_1 + \sum_{i=1}^q (\bar{\tau}_{i+1} - \bar{\tau}_i) \bar{\beta}_{i+1} \right], \tau \in [\bar{\tau}_0, \bar{\tau}_1].$$

Define

$$R(\tau, \tau^v, \tau^l) = \sigma^2 + \sum_{i=1}^{q+1} (\bar{\tau}_i - \bar{\tau}_{i-1}) \bar{\beta}_i' Q_{ff} \bar{\beta}_i - (\tau - \bar{\tau}_0) \bar{\beta}_1' Q \bar{\beta}_1 - (\bar{\tau}_{q+1} - \tau) \Gamma_2'(\tau) Q \Gamma_2(\tau)$$

on  $[\bar{\tau}_0, \bar{\tau}_1]$ , where  $Q = Q_{gf}' Q_{gg}^{-1} Q_{gf}$  is a symmetric positive definite matrix. Let  $T' = (\bar{\tau}_{q+1} - \bar{\tau}_0) T$ . By assumptions (A1) to (A8), and the use of triangle inequality, we have

$$\begin{aligned}& \sup_{\tau \in [\bar{\tau}_0, \bar{\tau}_1]} \left| \frac{1}{T'} RSS(\tau, \tau^v, \tau^l) - R(\tau, \tau^v, \tau^l) \right| \\ & \leq \left| \frac{1}{T'} U' U - \sigma^2 \right| + \left| \frac{1}{T'} \sum_{i=1}^{q+1} \bar{\beta}_i' F' I_i F \bar{\beta}_i - \sum_{i=1}^{q+1} (\tau_i - \tau_{i-1}) \bar{\beta}_i' Q_{ff} \bar{\beta}_i \right| \\ & \quad + \sup_{\tau \in [\bar{\tau}_0, \bar{\tau}_1]} \left| \frac{1}{T'} \left( \hat{\beta}_{a[\tau]}' G' I_a G \hat{\beta}_{a[\tau]} - 2 \hat{\beta}_{a[\tau]}' G' I_a F \bar{\beta}_1 \right) + (\tau - \bar{\tau}_0) \bar{\beta}_1' Q \bar{\beta}_1 \right| \\ & \quad + \sup_{\tau \in [\bar{\tau}_0, \bar{\tau}_1]} \left| \frac{1}{T'} \left( \hat{\beta}_{b[\tau]}' G' I_b G \hat{\beta}_{b[\tau]} - 2 \hat{\beta}_{b[\tau]}' G' (I_1 - I_a) F \bar{\beta}_1 - 2 \sum_{i=2}^{q+1} \hat{\beta}_{b[\tau]}' G' I_i F \bar{\beta}_i \right) \right. \\ & \quad \left. + (\bar{\tau}_{q+1} - \tau) \Gamma_2'(\tau) Q \Gamma_2(\tau) \right| \\ & \quad + \frac{2}{T'} \left| \sum_{i=1}^{q+1} U' I_i F \bar{\beta}_i \right| + \sup_{\tau \in [\bar{\tau}_0, \bar{\tau}_1]} \frac{2}{T'} \left| U' I_a G \hat{\beta}_{a[\tau]} \right| + \sup_{\tau \in [\bar{\tau}_0, \bar{\tau}_1]} \frac{2}{T'} \left| U' I_b G \hat{\beta}_{b[\tau]} \right| \\ & = o_p(1) \quad \text{since each individual term above is } o_p(1) .\end{aligned}$$

Thus,

$$\frac{1}{T'} RSS(\tau, \tau^v, \tau^l) \xrightarrow{p} R(\tau, \tau^v, \tau^l)$$

uniformly, where  $R(\tau, \tau^v, \tau^l)$  is a piecewise concave function of  $\tau$  defined on  $(\tau^v, \tau^l)$ .



Meanwhile,

$$\begin{aligned}
\frac{\partial R(\tau, \tau^v, \tau^l)}{\partial \tau} &= -\bar{\beta}'_1 Q \bar{\beta}_1 + \Gamma'_2(\tau) Q \Gamma_2(\tau) - 2(1-\tau) \Gamma'_2(\tau) Q \frac{\partial \Gamma_2(\tau)}{\partial \tau} \\
&= -\bar{\beta}'_1 Q \bar{\beta}_1 + \Gamma'_2(\tau) Q \Gamma_2(\tau) - 2\Gamma'_2(\tau) Q (\Gamma_2(\tau) - \bar{\beta}_1) \\
&= -(\Gamma_2(\tau) - \bar{\beta}_1)' Q (\Gamma_2(\tau) - \bar{\beta}_1) \\
&= -\left(\frac{1}{\bar{\tau}_{q+1} - \tau}\right)^2 L'_1 Q L_1 \leq 0,
\end{aligned}$$

where  $L_1 = \sum_{i=1}^q (\bar{\tau}_{i+1} - \bar{\tau}_i) \bar{\beta}_{i+1} - (\bar{\tau}_{q+1} - \bar{\tau}_1) \bar{\beta}_1$ .

$$\frac{\partial^2 R(\tau, \tau^v, \tau^l)}{\partial \tau^2} = -\frac{2}{(\bar{\tau}_{q+1} - \tau)^3} L'_1 Q L_1 \leq 0.$$

Therefore,  $\frac{1}{T} RSS(\tau, \tau^v, \tau^l)$  converges uniformly to a concave function of  $\tau$  for  $\tau \in [\bar{\tau}_0, \bar{\tau}_1]$ .

For  $\tau \in [\bar{\tau}_h, \bar{\tau}_{h+1})$ ,  $h = 1, 2, \dots, q-1$ , we have:

$$\begin{aligned}
I_a I_i &= I_i, 1 \leq i \leq h; \\
I_a I_{h+1} &= I_a - \sum_{i=1}^h I_i; \\
I_a I_i &= 0, h+1 < i \leq q+1; \\
I_b I_i &= 0, i \leq h; \\
I_b I_{h+1} &= \sum_{i=1}^{h+1} I_i - I_a; \\
I_b I_i &= I_i, h+1 < i \leq q+1.
\end{aligned}$$

Thus,

$$\begin{aligned}
\hat{\beta}_{a[\tau]} &= (G' I_a G)^{-1} G' I_a Y = \left(\frac{G' I_a G}{T'}\right)^{-1} \frac{1}{T'} G' I_a \left(\sum_{i=1}^{q+1} I_i F \bar{\beta}_i + U\right) \\
&= \left(\frac{G' I_a G}{T'}\right)^{-1} \frac{1}{T'} G' \left(\sum_{i=1}^h I_i F \bar{\beta}_i + I_a F \bar{\beta}_{h+1} - \sum_{i=1}^h I_i F \bar{\beta}_{h+1} + I_a U\right) \\
&\xrightarrow{p} [(\tau - \bar{\tau}_0) Q_{gg}]^{-1} Q_{gf} \left(\sum_{i=1}^h (\tau_i - \bar{\tau}_{i-1}) \bar{\beta}_i + (\tau - \tau_h) \bar{\beta}_{h+1}\right) \\
&= Q_{gg}^{-1} Q_{gf} \Gamma_1(\tau)
\end{aligned}$$

uniformly, where

$$\Gamma_1(\tau) = \frac{1}{\tau - \bar{\tau}_0} \left(\sum_{i=1}^h (\tau_i - \bar{\tau}_{i-1}) \bar{\beta}_i + (\tau - \tau_h) \bar{\beta}_{h+1}\right), \tau \in [\bar{\tau}_h, \bar{\tau}_{h+1}).$$

$$\begin{aligned}
\hat{\beta}_{b[\tau]} &= (G' I_b G)^{-1} G' I_b Y = (G' I_b G)^{-1} G' I_b \left(\sum_{i=1}^{q+1} I_i F \bar{\beta}_i + U\right) \\
&= \left(\frac{G' I_b G}{T'}\right)^{-1} \frac{1}{T'} G' \left[\left(\sum_{i=1}^{h+1} I_i - I_a\right) F \bar{\beta}_{h+1} + \sum_{i=h+2}^{q+1} I_i F \bar{\beta}_i + I_b U\right] \\
&\xrightarrow{p} [(\bar{\tau}_{q+1} - \tau) Q_{gg}]^{-1} Q_{gf} \left(\sum_{i=h+2}^{q+1} (\tau_i - \bar{\tau}_{i-1}) \bar{\beta}_i + (\bar{\tau}_{h+1} - \tau) \bar{\beta}_{h+1}\right) \\
&= Q_{gg}^{-1} Q_{gf} \Gamma_2(\tau)
\end{aligned}$$

uniformly, where

$$\Gamma_2(\tau) = \frac{1}{\bar{\tau}_{q+1} - \tau} \left( \sum_{i=h+2}^{q+1} (\tau_i - \bar{\tau}_{i-1}) \bar{\beta}_i + (\bar{\tau}_{h+1} - \tau) \bar{\beta}_{h+1} \right), \tau \in [\bar{\tau}_h, \bar{\tau}_{h+1}).$$

Define

$$\begin{aligned} R(\tau, \tau^v, \tau^l) &= \sigma^2 + \sum_{i=1}^{q+1} (\tau_i - \tau_{i-1}) \bar{\beta}'_i Q_{ff} \bar{\beta}_i \\ &\quad - (\tau - \bar{\tau}_0) \Gamma'_1(\tau) Q \Gamma_1(\tau) - (\bar{\tau}_{q+1} - \tau) \Gamma'_2(\tau) Q \Gamma_2(\tau) \end{aligned}$$

on  $[\bar{\tau}_h, \bar{\tau}_{h+1})$ , it follows that

$$\begin{aligned} & \sup_{\tau \in [\bar{\tau}_h, \bar{\tau}_{h+1})} \left| \frac{1}{T'} \text{RSS}(\tau, \tau^v, \tau^l) - R(\tau, \tau^v, \tau^l) \right| \\ & \leq \left| \frac{1}{T'} U' U - \sigma^2 \right| + \left| \frac{1}{T'} \sum_{i=1}^{q+1} \bar{\beta}'_i F' I_i F \bar{\beta}_i - \sum_{i=1}^{q+1} (\tau_i - \tau_{i-1}) \bar{\beta}'_i Q_{ff} \bar{\beta}_i \right| \\ & \quad + \sup_{\tau \in [\bar{\tau}_h, \bar{\tau}_{h+1})} \left| \frac{1}{T'} \begin{bmatrix} \hat{\beta}'_{a[\tau]} G' I_a G \hat{\beta}_{a[\tau]} - 2 \hat{\beta}'_{a[\tau]} G' \sum_{i=1}^h I_i F \bar{\beta}_i \\ -2 \hat{\beta}'_{a[\tau]} G' (I_a - \sum_{i=1}^h I_i) F \bar{\beta}_{h+1} \end{bmatrix} + (\tau - \bar{\tau}_0) \Gamma'_1(\tau) Q \Gamma_1(\tau) \right| \\ & \quad + \sup_{\tau \in [\bar{\tau}_h, \bar{\tau}_{h+1})} \left| \frac{1}{T'} \begin{bmatrix} \hat{\beta}'_{b[\tau]} G' I_b G \hat{\beta}_{b[\tau]} - 2 \hat{\beta}'_{b[\tau]} G' \left( \sum_{i=1}^{h+1} I_i - I_a \right) F \bar{\beta}_{h+1} - 2 \hat{\beta}'_{b[\tau]} G' \sum_{i=h+2}^{q+1} I_i F \bar{\beta}_i \\ + (\bar{\tau}_{q+1} - \tau) \Gamma'_2(\tau) Q \Gamma_2(\tau) \end{bmatrix} \right| \\ & \quad + \frac{2}{T'} \left| \sum_{i=1}^{q+1} U' I_i F \bar{\beta}_i \right| + \sup_{\tau \in [\bar{\tau}_h, \bar{\tau}_{h+1})} \frac{2}{T'} \left| U' I_a G \hat{\beta}_{a[\tau]} \right| + \sup_{\tau \in [\bar{\tau}_h, \bar{\tau}_{h+1})} \frac{2}{T'} \left| U' I_b G \hat{\beta}_{b[\tau]} \right| \\ & = o_p(1), \end{aligned}$$

meanwhile

$$\begin{aligned} \frac{\partial R(\tau, \tau^v, \tau^l)}{\partial \tau} &= - \left[ \begin{array}{c} \Gamma'_1(\tau) Q \Gamma_1(\tau) - \Gamma'_2(\tau) Q \Gamma_2(\tau) + \\ 2(\tau - \bar{\tau}_0) \Gamma'_1(\tau) Q \frac{\partial \Gamma_1(\tau)}{\partial \tau} + 2(\bar{\tau}_{q+1} - \tau) \Gamma'_2(\tau) Q \frac{\partial \Gamma_2(\tau)}{\partial \tau} \end{array} \right] \\ &= - \left[ \begin{array}{c} \Gamma'_1(\tau) Q \Gamma_1(\tau) - \Gamma'_2(\tau) Q \Gamma_2(\tau) + \\ 2\Gamma'_1(\tau) Q (\bar{\beta}_{h+1} - \Gamma_1(\tau)) + 2\Gamma'_2(\tau) Q (\Gamma_2(\tau) - \bar{\beta}_{h+1}) \end{array} \right] \\ &= \Gamma'_1(\tau) Q \Gamma_1(\tau) - \Gamma'_2(\tau) Q \Gamma_2(\tau) - 2\bar{\beta}'_{h+1} Q (\Gamma_1(\tau) - \Gamma_2(\tau)) \\ &= (\bar{\beta}_{h+1} - \Gamma_1(\tau))' Q (\bar{\beta}_{h+1} - \Gamma_1(\tau)) - (\bar{\beta}_{h+1} - \Gamma_2(\tau))' Q (\bar{\beta}_{h+1} - \Gamma_2(\tau)) \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 R(\tau, \tau^v, \tau^l)}{\partial \tau^2} &= 2\Gamma'_1(\tau) Q \frac{\partial \Gamma_1(\tau)}{\partial \tau} - 2\Gamma'_2(\tau) Q \frac{\partial \Gamma_2(\tau)}{\partial \tau} \\
&\quad - 2\bar{\beta}'_{h+1} Q \left[ \frac{1}{\tau - \bar{\tau}_0} (\bar{\beta}_{h+1} - \Gamma_1(\tau)) - \frac{1}{\bar{\tau}_{q+1} - \tau} (\Gamma_2(\tau) - \bar{\beta}_{h+1}) \right] \\
&= \frac{2}{\tau - \bar{\tau}_0} \Gamma'_1(\tau) Q (\bar{\beta}_{h+1} - \Gamma_1(\tau)) - \frac{2}{\bar{\tau}_{q+1} - \tau} \Gamma'_2(\tau) Q (\Gamma_2(\tau) - \bar{\beta}_{h+1}) \\
&\quad - 2\bar{\beta}'_{h+1} Q \bar{\beta}_{h+1} \left( \frac{1}{\tau - \bar{\tau}_0} + \frac{1}{\bar{\tau}_{q+1} - \tau} \right) + 2\bar{\beta}'_{h+1} Q \left( \frac{\Gamma_1(\tau)}{\tau - \bar{\tau}_0} + \frac{\Gamma_2(\tau)}{\bar{\tau}_{q+1} - \tau} \right) \\
&= -\frac{2}{\tau - \bar{\tau}_0} (\bar{\beta}'_{h+1} Q \bar{\beta}_{h+1} - 2\bar{\beta}'_{h+1} Q \Gamma_1(\tau) + \Gamma'_1(\tau) Q \Gamma_1(\tau)) \\
&\quad - \frac{2}{\bar{\tau}_{q+1} - \tau} (\bar{\beta}'_{h+1} Q \bar{\beta}_{h+1} - 2\bar{\beta}'_{h+1} Q \Gamma_2(\tau) + \Gamma'_2(\tau) Q \Gamma_2(\tau)) \\
&= -\frac{2}{\tau - \bar{\tau}_0} (\bar{\beta}_{h+1} - \Gamma_1(\tau))' Q (\bar{\beta}_{h+1} - \Gamma_1(\tau)) \\
&\quad - \frac{2}{\bar{\tau}_{q+1} - \tau} (\bar{\beta}_{h+1} - \Gamma_2(\tau))' Q (\bar{\beta}_{h+1} - \Gamma_2(\tau)) \\
&\leq 0.
\end{aligned}$$

Therefore,  $\frac{1}{T} RSS(\tau, \tau^v, \tau^l)$  converges uniformly to a concave function of  $\tau$  for  $\tau \in [\bar{\tau}_h, \bar{\tau}_{h+1})$ ,  $h = 1, 2, \dots, q-1$ .

For  $\tau \in [\bar{\tau}_q, \bar{\tau}_{q+1}]$ , we have

$$I_a I_i = I_i, i = 1, 2, \dots, q;$$

$$I_a I_{q+1} = I_a - \sum_{i=1}^q I_i;$$

$$I_b I_i = 0, i = 1, 2, \dots, q;$$

$$I_b I_{q+1} = I_b.$$

Using similar trick, we can prove that

$$\hat{\beta}_{a[\tau]} \xrightarrow{p} Q_{gg}^{-1} Q_{gf} \Gamma_1(\tau)$$

where

$$\Gamma_1(\tau) = \frac{1}{\tau - \bar{\tau}_0} \left[ \sum_{i=1}^q (\tau_i - \bar{\tau}_{i-1}) \bar{\beta}_i + (\tau - \bar{\tau}_q) \bar{\beta}_{q+1} \right], \tau \in [\tau_q, \tau_{q+1})$$

and

$$\hat{\beta}_{b[\tau]} \xrightarrow{p} Q_{gg}^{-1} Q_{gf} \bar{\beta}_{q+1}$$

Define

$$\begin{aligned}
R(\tau, \tau^v, \tau^l) &= \sigma^2 + \sum_{i=1}^{q+1} (\tau_i - \tau_{i-1}) \bar{\beta}'_i Q_{ff} \bar{\beta}_i \\
&\quad - (\tau - \bar{\tau}_0) \Gamma'_1(\tau) Q \Gamma_1(\tau) - (\bar{\tau}_{q+1} - \tau) \bar{\beta}'_{q+1} Q \bar{\beta}_{q+1}
\end{aligned}$$

on  $[\bar{\tau}_q, \bar{\tau}_{q+1}]$ , and it follows that

$$\sup_{\tau \in [\tau_q, \tau_{q+1}]} \left| \frac{1}{T'} RSS(\tau, \tau^v, \tau^l) - R(\tau, \tau^v, \tau^l) \right| = o_p(1).$$

Also,

$$\frac{\partial R(\tau, \tau^v, \tau^l)}{\partial \tau} = \frac{1}{\tau^2} L'_2 Q L_2 \geq 0,$$

where  $L_2 = \sum_{i=1}^q (\bar{\tau}_i - \bar{\tau}_{i-1}) \bar{\beta}_i - \bar{\tau}_q \bar{\beta}_{q+1}$ .

$$\frac{\partial^2 R(\tau, \tau^v, \tau^l)}{\partial \tau^2} = -\frac{2}{\tau^3} L'_2 Q L_2 \leq 0.$$

Combining the results above, we prove that  $\frac{1}{T}RSS(\tau, \tau^v, \tau^l)$  converges to an piecewise concave function  $R(\tau, \tau^v, \tau^l)$  on any segment  $(\tau^v T, \tau^l T)$  of  $[0, T]$ . With little modification, one can show that:

$$\frac{1}{T}RSS(\tau, 0, 1) \xrightarrow{P} R(\tau, 0, 1), \tau \in [\underline{\tau}, \bar{\tau}]$$

uniformly, where  $R(\tau, 0, 1)$  is piecewise concave function of  $\tau$  on  $(0, 1)$  and is defined as

$$R(\tau, 0, 1) = \sigma^2 + \sum_{i=1}^{p+1} (\tau_i - \tau_{i-1}) \beta'_i Q_{ff} \beta_i - \tau \Gamma'_1(\tau) Q \Gamma_1(\tau) - (1 - \tau) \Gamma'_2(\tau) Q \Gamma_2(\tau)$$

with  $\Gamma_1(\tau)$  and  $\Gamma_2(\tau)$  as defined in Section 3.  $\square$

### Proof of Theorem 3.3: Rewrite

$$f(x_t) = Q_{fg} Q_{gg}^{-1} g(x_t) + e_t,$$

where  $Q_{gg} = E(g_t g'_t)$ . By definition,  $E(e_t g'_t) = 0$ , where  $g_t = g(x_t)$ . In matrix notation,

$$F = G Q_{gg}^{-1} Q_{gf} + e,$$

where

$$e = F - G Q_{gg}^{-1} Q_{gf}.$$

Thus,

$$Y = G\delta + e\beta + \varepsilon = G\delta + v,$$

where

$$\delta = Q_{gg}^{-1} Q_{gf} \beta,$$

$$v = e\beta + \varepsilon.$$

Under the null, we have,

$$\hat{\beta} = (G'G)^{-1} G'Y = \delta + (G'G)^{-1} G'v.$$

Thus,

$$\hat{\beta} - \delta = (G'G)^{-1} G' (e\beta + \varepsilon).$$

While  $e_t$  is uncorrelated with  $g_t$ ,  $(\beta' e_t)^2$  is not necessarily uncorrelated with  $g_t g'_t$  under measurement error or misspecification. This means that the equality  $E[(\beta' e_t)^2 g_t g'_t] = E(\beta' e_t)^2 E(g_t g'_t)$  does not hold in general. Thus, we need to use the HAC robust Wald test. The whole sample estimator  $\hat{\beta}$  has the limiting distribution

$$\sqrt{T}(\hat{\beta} - \delta) \xrightarrow{d} N(0, \Omega),$$

where  $\Omega = Q_{gg}^{-1} V_L Q_{gg}^{-1}$ ,  $V_L$  the long-run covariance matrix for  $g_t v_t = g_t e'_t \beta + g_t \varepsilon_t$ .

For the first subsample, we have

$$\sqrt{T}(\hat{\beta}_{a[\tau]} - \delta) \xrightarrow{d} \frac{1}{\tau} \Omega^{1/2} B(\tau),$$

where  $B(\tau)$  is a vector of independent of Brownian motions.

Similarly, for the second subsample estimator, we have

$$\sqrt{T}(\hat{\beta}_{b[\tau]} - \delta) \xrightarrow{d} \frac{1}{1-\tau} \Omega^{1/2} [B(1) - B(\tau)].$$

Combining the two results, we have

$$\sqrt{T}(\hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]}) \xrightarrow{d} \frac{1}{\tau(1-\tau)}[B(\tau) - \tau B(1)].$$

Assuming no serial correlation for simplicity, the White (1980) estimator for the covariance matrix  $V_L$  is

$$\hat{V}_L = \frac{1}{T} \sum_{t=1}^T g_t g_t' \hat{v}_t^2.$$

The Wald test

$$T\tau(1-\tau)(\hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]})' \hat{\Omega}^{-1} (\hat{\beta}_{a[\tau]} - \hat{\beta}_{b[\tau]}) \xrightarrow{d} \frac{\|B(\tau) - \tau B(1)\|^2}{\tau(1-\tau)}$$

where  $\hat{\Omega} = \hat{Q}_{gg}^{-1} \hat{V}_L \hat{Q}_{gg}^{-1}$  and  $\hat{Q}_{gg} = \frac{G'G}{T}$ .