Inference on a Structural Break in Trend with Fractionally Integrated Errors

Seongyeon Chang [*] **Pierre Perron** [†]

Boston University **Boston University**

November 12, 2013

Abstract

Perron and Zhu (2005) established the consistency, rate of convergence, and the limiting distributions of parameter estimates in a linear time trend with a change in slope with or without a concurrent change in level. They considered the dichotomous cases whereby the errors are short-memory, $I(0)$, or have an autoregressive unit root, $I(1)$. We extend their analysis to cover the more general case of fractionally integrated errors for values of d^* in the interval $(-0.5, 1.5)$ excluding the boundary case 0.5. Our theoretical results uncover some interesting features. For example, when a concurrent level shift is allowed, the rate of convergence of the estimate of the break date is the same for all values of d^* in the interval $(-0.5, 0.5)$. This feature is linked to the contamination induced by allowing a level shift, previously discussed by Perron and Zhu (2005). In all other cases, the rate of convergence is monotonically decreasing as d^* increases. We also provide results about the so-called spurious break issue. Simulation experiments are provided to illustrate some of the theoretical results in the paper.

JEL Classification: C22.

Keywords: long-memory, segmented trend, structural change, spurious break, linear trend.

[*]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA 02215 (seongch@bu.edu)
[†]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA 02215 (perron@bu.edu)

1 Introduction

Economic relationships are often subject to structural changes. Hence, testing for a structural break and estimating the break date have been important topics in both economics and statistics; see Perron (2006) for a review. To test for a structural break, or instability of the parameters, important contributions include Andrews (1993) and Andrews and Ploberger (1994). Bai (1994, 1997) showed that the break date can be estimated consistently by minimizing the sum of squared residuals (SSR) from the unrestricted model and derived the limiting distribution of the estimate of the break date, which can be applied to construct confidence intervals (CIs) for the true break date. Bai and Perron (1998, 2003) considered statistical inference related to multiple structural changes under general conditions.

In the literature, most of the work assumed that the regressors and the errors are stationary. Structural breaks in trend regressors and non-stationary processes are also important from a practical perspective. Perron (1989) showed that the Dickey and Fuller (1979) type unit root test is biased in favor of a non-rejection of the unit root null hypothesis when the process is trend stationary with a structural break in slope. With respect to the problem of estimating the break date of the change in the slope of a linear trend with or without a concurrent level shift, Perron and Zhu (2005) (PZ, henceforth) analyzed the consistency, rate of convergence and the limiting distributions of the parameter estimates when the errors are either short-memory, $I(0)$, or have an autoregressive unit root, $I(1)$. We extend their analysis to cover the more general case of fractionally integrated errors for values of d^* in the interval $(-0.5,\ 1.5)$ excluding the boundary case 0.5. Our theoretical results uncover some interesting features. For example, when a concurrent level shift is allowed, the rate of convergence of the estimate of the break date is the same for all values of d^* in the interval $(-0.5, 0.5)$. This feature is linked to the contamination induced by allowing a level shift, previously discussed by Perron and Zhu (2005). In all other cases, the rate of convergence is monotonically decreasing as d^* increases. We also provide results about the so-called spurious break issue and show that it cannot occur for all values of d^* in the interval $(-0.5, 0.5)$. Simulation experiments are provided to illustrate the theoretical results in the paper.

Work related to changes in trend include the following. Feder (1975) considered estimating the joint points of polynomial type segmented regressions. Bai (1997) and Bai and Perron (1998) provided inference results with trending regressors. In order to obtain the limiting distribution, the trending regressors are assumed to be a function of t/T, say $g(t/T)$, with T the sample size. Deng and Perron (2006) analyzed the consequences of specifying the trend function in scaled form when a structural break is involved. Bai et al. (1998) analyzed the limiting distribution of the estimated break date for non-stationary type series with a change in slope. Chu and White (1992) suggested a testing procedure for a change in a trend function with stationary errors. Perron (1991) and Vogelsang (1997) considered testing a structural break in trend when the errors are either stationary or have a unit root. Vogelsang (1999) devised a test whose limiting distribution does not change depending on whether the noise component is stationary or integrated. Recently, Perron and Yabu (2009) considered testing for structural changes in the trend function of a time series without any prior knowledge about whether the errors are stationary or integrated. Their testing procedure adopts a quasi-feasible generalized least squares approach that uses a super-efficient estimate of the sum of the autoregressive parameters α when $\alpha = 1$. Harvey et al. (2009) proposed a generalized least squares (GLS)-based trend break test that is asymptotically size robust with $I(0)$ and $I(1)$ errors. The results of PZ and Perron and Yabu (2009) have been used in Kim and Perron (2009) to provide unit root tests with improved power that allow for a change in the trend function under both the null and alternative hypotheses.

Fractionally integrated processes have been popular in the economics and statistics literature, in particular following the introduction of the ARFIMA processes by Granger and Joyeux (1980) and Hosking (1981). Kuan and Hsu (1998) considered a change in mean model and showed the consistency and the rate of convergence of the least square estimate of the break date when the errors are fractionally integrated; see also Lavielle and Moulines (2000). They found that the convergence rate depends on the order of integration d^*. Moreover, when no such change in mean is present, the estimate of the break date obtained by minimizing the sum of squared residuals supports a spurious break date when $d^* \in (0, 0.5)$. Hsu and Kuan (2008) showed that the least square estimate of the break date in a mean change model is not consistent when the errors are fractionally integrated with $d^* \in (0.5, 1.5)$, and that the spurious feature also occurs. Gil-Alana (2008) executed a set of Monte Carlo simulations to confirm that both the order of fractional integration and the break date can be estimated simultaneously by minimizing the SSR considering all possible grids on d^* and break dates T_1. More recently, Iacone et al. (2013) provided a sup-Wald type test for a structural change in trend when the order of integration in the errors is unknown a priori.

The structure of the paper is as follows. In section 2, we review fractionally integrated processes, fractional Brownian motion and useful related functional central limit theorems. Section 3 presents the models, the assumptions and a key inequality used throughout the proofs. Section 4 provides the main contributions related to the limit properties of the estimates: consistency (Section 4.1), rate of convergence (Section 4.2), limit distributions of the estimate of the break date (Section 4.3) and limit distributions of the estimates of the other parameters (Section 4.4). The problem of the possibility of a spurious break is discussed in Section 5. Section 6 provides brief concluding remarks and an appendix contains all technical derivations.

2 Fractionally Integrated Processes and Functional Central Limit Theorem

In this section, we briefly define fractionally integrated processes and review results to be used in subsequent developments. We follow the notation of Marinucci and Robinson (1999) and Wang et al. (2003).

Definition 1 *An autoregressive fractionally integrated moving average (ARFIMA) process u_t is defined as*

$$(1-L)^{d+m}u_t = \xi_t \quad \text{and} \quad \xi_t = \psi(L)\varepsilon_t \tag{1}$$

for $t = 1, 2, \ldots$, where $m \geq 0$ is an integer and $d \in (-0.5, 0.5)$; L is the lag operator such that $Lu_t = u_{t-1}$ and ε_t are independently and identically distributed (i.i.d.) random variables with zero mean and finite variance.

Using this notation, note that the order of integration is $d^* = m + d$. Wang et al. (2003) derived the invariance principle for $m \geq 0$ which includes the non-stationary cases. We summarize their results insofar as they will be relevant for subsequent derivations.

Lemma 1 *(Wang et al., 2003, Theorem 2.1): Let u_t satisfy (1) with $m = 0$ and let $\{\psi_j; j \geq 0\}$ satisfy*

$$\sum_{j=0}^{\infty} |\psi_j| < \infty \quad \text{and} \quad b_\psi \equiv \sum_{j=0}^{\infty} \psi_j \neq 0 \tag{2}$$

Assume that $E\epsilon_0^2 < \infty$. Then, for $0 \leq d < 0.5$,

$$\frac{1}{\kappa(d)T^{1/2+d}} \sum_{j=1}^{[Tr]} u_j \Rightarrow W_d(r), \quad r \in [0,1], \tag{3}$$

where

$$\kappa^2(d) = \frac{b_\psi^2 \Gamma(1-2d) E\epsilon_0^2}{(1+2d)\Gamma(1+d)\Gamma(1-d)} \tag{4}$$

and $W_d(r)$ is a Type I fractional Brownian motion [1] *on $D[0,1]$. Also, if $E|\epsilon_0|^{(2+c)/(1+2d)} < \infty$, where $c > 0$, then* (3) *still holds for $d \in (-0.5, 0)$.*

Lemma 2 *(Wang et al., 2003, Theorem 2.2): Let u_t satisfy* (1) *with $m = 0$ and let $\{\psi_j; j \geq 0\}$ satisfy*

$$\sum_{j=0}^{\infty} j^{1/2-d}|\psi_j| < \infty \quad \textit{and} \quad b_\psi \equiv \sum_{j=0}^{\infty} \psi_j \neq 0. \tag{5}$$

Assume that $E|\epsilon_0|^{\max\{2,2/(1+2d)\}} < \infty$. Then, (3) *holds for $d \in (-0.5, 0.5)$.*

Corollary 1 *(Wang et al., 2003, Corollary 2.1): Let u_j satisfy* (1) *with $m = 0$. If $\mathbf{E}|\epsilon_0|^{\max\{2,2/(1+2d)\}} < \infty$, then* (3) *follows with $b_\psi = \theta(1)/\phi(1)$ for $d \in (-0.5, 0.5)$.*

In order to consider general non-stationary fractionally integrated processes, two additional conditions are required.

- **Condition 1**: $\{\psi_j; j \geq 0\}$ satisfy (2), and $E|\varepsilon_0|^p < \infty$ where $p = 2$, for $d \in [0, 0.5)$; $p = (2+c)/(1+2d) < \infty$, $c > 0$ for $d \in (-0.5, 0)$.
- **Condition 2**: $\{\psi_j; j \geq 0\}$ satisfy (5), and $E|\varepsilon_0|^{\max\{2,2/(1+2d)\}} < \infty$.

Lemma 3 *(Wang et al. , 2003, Theorem 3.1): Let u_j satisfy* (1) *with $m \geq 1$. Let Conditions 1 and 2 hold. Then, for $0 \leq r \leq 1$,*

$$\frac{1}{\kappa(d)T^{-1/2+d+m}} u_{[Tr]} \Rightarrow W_{d,m}(r),$$

$$\frac{1}{\kappa(d)T^{1/2+d+m}} \sum_{j=1}^{[Tr]} u_j \Rightarrow \int_0^r \int_0^{r_m} \cdots \int_0^{r_2} W_d(r_1) dr_1 dr_2 \ldots dr_m,$$

$$\frac{1}{\kappa^2(d)T^{2(d+m)}} \sum_{j=1}^{[Tr]} u_j^2 \Rightarrow \int_0^r [W_{d,m}(s)]^2 ds$$

[1] Marinucci and Robinson (1999) defined type I fractional Brownian motions with $d \in (-0.5, 0.5)$ on $D[0,1]$ as follows:

$$W_d(t) = \frac{1}{A(d)} \int_{-\infty}^{0} [(t-s)^d - (-s)^d] dW(d) + \int_0^t (t-s)^d dW(s),$$

where $W(s)$ is a standard Brownian motion and

$$A(d) = \left(\frac{1}{2d+1} + \int_0^\infty [(1+s)^d - s^d]^2 ds \right)^{1/2}.$$

They explained the difference between Type I and Type II fractional Brownian motion and showed how those two types are related.

where

$$W_{d,m}(r) = \begin{cases} W_d(r) & \text{if } m = 1 \\ \int_0^r \int_0^{r_{m-1}} \cdots \int_0^{r_2} W_d(r_1) dr_1 dr_2 \cdots dr_{m-1} & \text{if } m \geq 2. \end{cases}$$

We shall be interested in the case $m = 1$, in which case

$$\frac{1}{\kappa(d) T^{1/2+d}} \Rightarrow W_d(r), \tag{6}$$

$$\frac{1}{\kappa(d) T^{1/2+d+1}} \sum_{j=1}^{[Tr]} u_j \Rightarrow \int_0^r W_d(s) ds,$$

$$\frac{1}{\kappa^2(d) T^{2(d+1)}} \sum_{j=1}^{[Tr]} u_j^2 \Rightarrow \int_0^r [W_d(s)]^2 ds.$$

3 The Models and Assumptions

The series of interest, y_t, is assumed to consist of some systematic part f_t and a random component u_t, namely,

$$y_t = f_t + u_t.$$

For the noise component u_t, we assume $Eu_t = 0$ and that the following two assumptions hold.

- **Assumption A1:** For $d^* \in (-0.5, 0.5) \cup (0.5, 1.5)$, $u_t = \Delta^{-d^*} \xi_t \mathbf{1}_{t>0}$ where $\Delta^{d^*} \equiv (1-L)^{d^*}$, with L being a lag operator such that $L\xi_t = \xi_{t-1}$ and $\Delta^{-d^*} = \sum_{j=0}^{\infty} \pi_j L^j$ with $\pi_j \equiv \Gamma(j + d^*)/[\Gamma(d^*)\Gamma(j+1)]$, where $\Gamma(\cdot)$ denotes the Gamma function.

- **Assumption A2:** $\xi_t \sim I(0)$. More specifically, ξ_t is such that $T^{-1/2} \sum_{t=1}^{[Tr]} \xi_t \Rightarrow b_\psi W(r)$ where $b_\psi^2 = \lim_{T\to\infty} T^{-1} \mathbf{E}(\sum_{t=1}^{T} \xi_t)^2$ exists and is strictly positive. Here $\Rightarrow$ denotes weak convergence in distribution (under the Skorohod topology) and $W(\cdot)$ is the standard Wiener process.

Remark 1 *There are many sets of sufficient conditions to guarantee that the invariance principle in Assumption A2 holds. For example, it holds when ξ_t is a linear process such that $\xi_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$ with $\sum_{j=0}^{\infty} j|c_j| < \infty$, where $\{\varepsilon_t, \mathcal{F}_{t-1}\}$ is a martingale difference sequence with $\mathcal{F}_{t-1}$ the filtration to which ε_t is adapted.*

For the systematic part f_t, we consider two cases studied in PZ with $I(0)$ and $I(1)$ errors. The first, labelled Model I, specifies that f_t is a first-order linear trend with a single change in slope. In this case, the trend is joined at the time of break and there is no concurrent level shift. The second, labelled Model II, specifies that f_t is a first-order linear trend with a concurrent break in both intercept and slope. Let $\lambda = T_1/T$ denote the break fraction with a postulated break date T_1.

- **Model I: Joint Broken Trend with Fractionally Integrated Errors**: The deterministic component f_t is specified as

$$f_t = \mu_1 + \beta_1 t + \beta_b B_t, \tag{7}$$

where B_t is a dummy variable for the slope change defined by

$$B_t = \begin{cases} 0 & \text{if } t \leq T_1, \\ t - T_1 & \text{if } t > T_1. \end{cases}$$

Hence, the slope coefficient changes from β_1 to $\beta_1 + \beta_b$ at time T_1. Note that the trend function is continuous at the time point T_1, hence the labelling of a "joint broken trend".

- **Model II: Local Disjoint Broken Trend with Fractionally Integrated Errors**: The deterministic component f_t is specified by

$$f_t = \mu_1 + \beta_1 t + \mu_b C_t + \beta_b B_t, \tag{8}$$

where C_t is a dummy variable for the level shift defined by

$$C_t = \begin{cases} 0 & \text{if } t \leq T_1, \\ 1 & \text{if } t > T_1. \end{cases}$$

At the break date T_1, there are concurrent slope and level shifts. The magnitude of the level shift is μ_b, which is asymptotically negligible compared to the level of the series $\mu_1 + \beta_1 T_1$, hence the labelling of a "local disjoint broken trend".

Throughout, we assume that there is at least a change in slope as stated in the following assumption. Let the true break date be denoted by T_1^0 and let the break fraction be $\lambda_0 = T_1^0/T$.

- **Assumption A3:** $\beta_b \neq 0$ and $\lambda_0 \in (0, 1)$.

This assumption is required to ensure that there is a break in slope and that the pre and post break samples are asymptotically large enough to obtain consistent estimates of the unknown coefficients. This is a standard assumption needed to derive any useful asymptotic result. By construction, the true break date T_1^0 increases in the sample size T.

In matrix notation, the Data Generating Processes can be specified as follows

$$Y = X_{T_1^0}\gamma + U \tag{9}$$

where $Y = [y_1, \ldots, y_T]'$, $U = [u_1, \ldots, u_T]'$, $X_k = [x(k)_1, \ldots, x(k)_T]'$, with $x(k)_t' = [1 \quad t \quad B_t]$ and $\gamma = [\mu_1 \quad \beta_1 \quad \beta_b]'$, for Model I, while for Model II, $x(k)_t' = [1 \quad t \quad C_t \quad B_t]$ and $\gamma = [\mu_1 \quad \beta_1 \quad \mu_b \quad \beta_b]'$. Note that the matrix X_{T_1} depends on the candidate break date T_1, while $X_{T_1^0}$ depends on the true break date T_1^0. The parameters are estimated using a global least-squares criterion. The estimate of the break date is

$$\hat{T}_1 = \arg\min_{T_1} Y'(I - P_{T_1})Y$$

where P_{T_1} is the matrix that projects on the range space of X_{T_1}, i.e., $P_{T_1} = X_{T_1}(X_{T_1}'X_{T_1})^{-1}X_{T_1}'$. With $X_{\hat{T}_1}$ constructed using the estimate of the break date $\hat{T}_1$, the OLS estimate of γ is

$$\hat{\gamma} = (X_{\hat{T}_1}'X_{\hat{T}_1})^{-1}X_{\hat{T}_1}'Y$$

and the resulting sum of squared residuals is, for an estimated break fraction $\hat{\lambda} = \hat{T}_1/T$,

$$S(\hat{\lambda}) = \sum_{t=1}^{T}\hat{u}_t^2 = \sum_{t=1}^{T}\left(y_t - x(\hat{T}_1)_t'\hat{\gamma}\right)^2 = Y'(I - P_{\hat{T}_1})Y$$

where $P_{\hat{T}_1}$ is the projection matrix associated with $X_{\hat{T}_1}$. The limiting distributions of $\hat{\lambda} - \lambda_0$ and $\hat{\gamma} - \gamma_0$ have been derived by PZ for the dichotomous case with either $I(0)$ and $I(1)$ errors.

In what follows, our aim is the following. First, we show that the break fraction λ_0 can be estimated consistently by minimizing the sum of squared residuals when the errors are fractionally integrated. Second, we derive the limit distributions of the estimates. Third, we show that a spurious break phenomenon can occur even in the case of a break in a linear time trend when the errors are fractionally integrated.

3.1 A Key Inequality

As in PZ, a key inequality will play a crucial role in proving the limit results. First, by construction, we have for all T,

$$S(\hat{\lambda}) \leq S(\lambda_0)$$

or equivalently,

$$Y'(I - P_{\hat{T}_1})Y \leq Y'(I - P_{T_1^0})Y.$$

Using (9), this inequality can be rewritten as

$$Y'(P_{T_1^0} - P_{\hat{T}_1})Y \leq 0$$

or equivalently,

$$\begin{aligned}
&(\gamma_0' X_{T_1^0}' + U')(P_{\hat{T}_1^0} - P_{\hat{T}_1})(X_{T_1^0}\gamma_0 + U) \\
&= \gamma_0' X_{T_1^0}'(P_{T_1^0} - P_{\hat{T}_1})X_{T_1^0}\gamma_0 + 2\gamma_0' X_{T_1^0}'(P_{T_1^0} - P_{\hat{T}_1})U + U'(P_{T_1^0} - P_{\hat{T}_1})U \\
&= \gamma_0'(X_{T_1^0} - X_{\hat{T}_1})'(I - P_{\hat{T}_1})(X_{T_1^0} - X_{\hat{T}_1})\gamma_0 \\
&\quad + 2\gamma_0'(X_{T_1^0} - X_{\hat{T}_1})'(I - P_{\hat{T}_1})U + U'(P_{T_1^0} - P_{\hat{T}_1})U \\
&\equiv (XX) + 2(XU) + (UU) \leq 0
\end{aligned} \tag{10}$$

where we have made use of the fact that $X_{T_1^0}' P_{T_1^0} = X_{T_1^0}'$ and $X_{T_1^0}'(I - P_{T_1^0}) = 0$. Moreover, it is straightforward to show that

$$\begin{aligned}
\arg\min{}_{T_1}[SSR(T_1)] &= \arg\min{}_{T_1}[SSR(T_1) - SSR(T_1^0)] \\
&= \arg\min{}_{T_1}[\gamma_0'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})(X_{T_1^0} - X_{T_1})\gamma_0 \\
&\qquad + 2\gamma_0'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})U + U'(P_{T_1^0} - P_{T_1})U].
\end{aligned}$$

We will make use of this result later to derive the rate of convergence of $\hat{\lambda} = \hat{T}_1/T$.

4 Asymptotic Results

We consider in turn the consistency, rate of convergence and limit distribution of the estimates, concentrating on the estimate of the break date.

4.1 Consistency

We show that $\hat{\lambda}$ is a consistent estimate of λ_0 when the errors are fractionally integrated with parameter $d^* \in (-0.5, 0.5) \cup (0.5, 1.5)$. The idea behind the proof is the following. Unless $\hat{\lambda} \rightarrow_p \lambda_0$, the first (non-negative) term in (10) would asymptotically dominate the others. It means that the key inequality does not hold if $\hat{\lambda}$ does not converge to λ_0, which leads to the desired contradiction. We start with the following theorem.

Theorem 1 *Define*

$$
\begin{aligned}
(XX) &= \gamma_0'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})(X_{T_1^0} - X_{T_1})\gamma_0 \\
(XU) &= \gamma_0'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})U \\
(UU) &= U'(P_{T_1^0} - P_{T_1})U.
\end{aligned}
$$

Under the Assumptions A1-A3, the following results hold uniformly over all generic $T_1 \in [\pi T, (1-\pi)T]$ for some arbitrary small π such that $\lambda_0 \in [\pi, 1-\pi]$. First, for $d \in (-0.5, 0.5)$, we have for Model I: i) if $m = 0$:

$$
\begin{aligned}
(XX) &= |T_1 - T_1^0|^2 O(T) \\
(XU) &= |T_1 - T_1^0| O_p(T^{1/2+d}) \\
(UU) &= |T_1 - T_1^0| O_p(T^{-1+2d}),
\end{aligned}
$$

ii) if $m = 1$:

$$
\begin{aligned}
(XX) &= |T_1 - T_1^0|^2 O(T) \\
(XU) &= |T_1 - T_1^0| O_p(T^{3/2+d}) \\
(UU) &= |T_1 - T_1^0| O_p(T^{1+2d}).
\end{aligned}
$$

For Model II: i) if $m = 0$:

$$
\begin{aligned}
(XX) &= |T_1 - T_1^0|^3 O(1) \\
(XU) &= |T_1 - T_1^0|^{3/2+d} O_p(1) \\
(UU) &= |T_1 - T_1^0|^{1/2+d} O_p(T^{-1/2+d}).
\end{aligned}
$$

ii) if $m = 1$

$$
\begin{aligned}
(XX) &= |T_1 - T_1^0|^3 O(1) \\
(XU) &= |T_1 - T_1^0|^2 O_p(T^{1/2+d}) \\
(UU) &= |T_1 - T_1^0| O_p(T^{1+2d}).
\end{aligned}
$$

Note that (XX) is always positive because it is quadratic and $(I - P_{T_1})$ is positive semi-definite. Given the results in Theorem 1, unless $\hat{\lambda} \to_p \lambda_0$, then $(XX) = O(T^3)$, $(XU) = O_p(T^{3/2+d})$, and $(UU) = O_p(T^{2d})$ with $m = 0$. Similarly, $(XX) = O(T^3)$, $(XU) = O_p(T^{5/2+d})$, and $(UU) = O_p(T^{2+2d})$ with $m = 1$. Hence, for large T and $d \in (-0.5, 0.5)$, with some probability, the positive term (XX) dominates the other two terms (XU) and (UU) such that inequality (10) will not hold with probability 1. Hence, we have a contradiction since the inequality (10) holds by construction. Therefore, we can conclude that $\hat{\lambda} \to_p \lambda_0$, as stated in the following theorem.

Theorem 2 *Under Assumptions A1-A3, in Model I-II,* $\hat{\lambda} \xrightarrow{p} \lambda_0$, $\forall d^* \in (-0.5, 0.5) \cup (0.5, 1.5)$.

4.2 Rate of Convergence

The following theorem shows that the rate of the convergence of the estimate of the break fraction, $\hat{\lambda}$, depends on the order of fractional integration d^*. It also differs across the two models being faster with no concurrent level shift.

Theorem 3 *Under Assumptions A1-A3, for every* $d \in (-0.5, 0.5)$*: For Model I:*

$$\hat{\lambda} - \lambda_0 = \begin{cases} O_p(T^{-3/2+d}) & \text{if } m = 0 \\ O_p(T^{-1/2+d}) & \text{if } m = 1, \end{cases}$$

For Model II:

$$\hat{\lambda} - \lambda_0 = \begin{cases} O_p(T^{-1}) & \text{if } m = 0 \\ O_p(T^{-1/2+d}) & \text{if } m = 1. \end{cases}$$

Theorem 3 implies that the rate of convergence is slower when allowing for a concurrent level shift, even if none is present, when $d^* \in (-0.5, 0.5)$. It is, however, the same when $d^* \in (0.5, 1.5)$. These results accord with those from PZ who considered $I(0)$ and $I(1)$ processes. For Model I and II with $I(1)$ process, $\hat{\lambda} - \lambda_0 = O_p(T^{-1/2})$. On the other hand, for Model I with $I(0)$ process, $\hat{\lambda} - \lambda_0 = O_p(T^{-3/2})$ and for Model II with $I(0)$ process, $\hat{\lambda} - \lambda_0 = O_p(T^{-1})$. PZ (2005) presented an intuitive explanation for the change in convergence rate induced by introducing a level shift. Briefly, a random deviation from a deterministic trend function is subject to be captured as if it were a level shift. Hence, it can have an effect on the precision of the estimate.

The results show that the rate of convergence is linearly decreasing as d^* increases for all models except Model II for $d^* \in (-0.5, 0.5)$. The result for this latter case is quite interesting as the rate of convergence is the same for all $d^* \in (-0.5, 0.5)$. The explanation for this feature is again related to the contamination induced by allowing a concurrent level shift, which implies added noise. If the process is stationary, $d^* \in (-0.5, 0.5)$, this added noise dominates and renders the rate of convergence invariant to d^*. If the process is non-stationary, $d^* \in (0.5, 1.5)$, the noise is small compared to the signal and we are back essentially to the case with no concurrent level shift.

4.3 The Limiting distribution of the estimate of the break date

Given results about the convergence and the rate of convergence of the estimate of the break fraction $\hat{\lambda}$, we can now consider its limiting distribution. The results are stated in the following Theorem.

Theorem 4 *Under Assumptions A1-A3, we have for every $d \in (-0.5, 0.5)$: 1) For Model I: a) if $m = 0$,*

$$T^{3/2-d}(\hat{\lambda} - \lambda_0) \Rightarrow -\frac{4\kappa(d)\zeta}{\lambda_0(1-\lambda_0)\beta_b},$$

b) if $m = 1$,

$$T^{1/2-d}(\hat{\lambda} - \lambda_0) \Rightarrow -\frac{4\kappa(d)\int_{\lambda_0}^1 W_d^*(r)dr}{\lambda_0(1-\lambda_0)\beta_b}$$

where $\kappa(d)$ is defined by (4),

$$\begin{aligned}\zeta &= \int_{\lambda_0}^1 dW_d(r) + \frac{1-\lambda_0}{2}\int_0^1 dW_d(r) - \frac{3(1-\lambda_0)}{2\lambda_0}\int_0^1 r dW_d(r) \\ &\quad - \frac{3(2\lambda_0 - 1)}{2\lambda_0(1-\lambda_0)}\int_{\lambda_0}^1 (r-\lambda_0) dW_d(r),\end{aligned}$$

and

$$\begin{aligned}\int_{\lambda_0}^1 W_d^*(r)dr &= \int_{\lambda_0}^1 W_d(r)dr + \frac{1-\lambda_0}{2}\int_0^1 W_d(r)dr - \frac{3(1-\lambda_0)}{2\lambda_0}\int_0^1 rW_d(r)dr \\ &\quad - \frac{3(2\lambda_0-1)}{2\lambda_0(1-\lambda_0)}\int_{\lambda_0}^1 (r-\lambda_0)W_d(r)dr.\end{aligned}$$

2) For Model II: a) if $m = 0$, define a stochastic process $S^*(m)$ on the set of integers as follows: $S^*(0) = 0$, $S^*(m) = S_1(m)$ for $m < 0$ and $S^*(m) = S_2(m)$ for $m > 0$, with

$$\begin{aligned}S_1(m) &= \sum_{k=m+1}^{0}(\mu_b + \beta_b k)^2 - 2\sum_{k=m+1}^{0}(\mu_b + \beta_b k)u_k, \quad m = -1, -2, \ldots, \\ S_2(m) &= \sum_{k=1}^{m}(\mu_b + \beta_b k)^2 + 2\sum_{k=1}^{m}(\mu_b + \beta_b k)u_k, \quad m = 1, 2, \ldots.\end{aligned}$$

If u_t is strictly stationary with a continuous distribution, S^* is a two-sided random walk with drift, and $T^{1-d}(\hat{\lambda}-\lambda) \Rightarrow \arg\min_m S^*(m)$. b) if $m = 1$, define

$$\begin{aligned}
\xi_1 &= \left(\int_0^1 W_d(r)dr, \int_0^1 rW_d(r)dr, \int_{\lambda_0}^1 W_d(r)dr, \int_{\lambda_0}^1 (r-\lambda_0)W_d(r)dr\right)', \\
\xi_2 &= \left(0, 0, W_d(\lambda_0), \int_{\lambda_0}^1 W_d(r)dr\right)', \\
\xi_3 &= \int_0^{\lambda_0} [(3r^2 - 2r\lambda_0)/(\lambda_0)^2] dW_d(r), \\
\xi_4 &= \int_{\lambda_0}^1 [(r-1)(3r - 2\lambda_0 - 1)/(1-\lambda)^2] dW_d(r),
\end{aligned}$$

$$\Omega_1 = \begin{bmatrix} \frac{4}{\lambda_0} & -\frac{6}{\lambda_0^2} & \frac{2}{\lambda_0} & \frac{6}{\lambda_0^2} \\ -\frac{6}{\lambda_0^2} & \frac{12}{\lambda_0^3} & -\frac{6}{\lambda_0^2} & -\frac{12}{\lambda_0^3} \\ \frac{2}{\lambda_0} & -\frac{6}{\lambda_0^2} & \frac{4}{\lambda_0(1-\lambda_0)} & \frac{6(1-2\lambda_0)}{\lambda_0^2(1-\lambda_0)^2} \\ \frac{6}{\lambda_0^2} & -\frac{12}{\lambda_0^3} & \frac{6(1-2\lambda_0)}{\lambda_0^2(1-\lambda_0)^2} & \frac{12(3\lambda_0^2-3\lambda_0+1)}{\lambda_0^3(1-\lambda_0)^3} \end{bmatrix}$$

$$\Omega_2 = \begin{bmatrix} -\frac{4}{\lambda_0^2} & \frac{12}{\lambda_0^3} & -\frac{2}{\lambda_0^2} & -\frac{12}{\lambda_0^3} \\ \frac{12}{\lambda_0^3} & -\frac{36}{\lambda_0^4} & \frac{12}{\lambda_0^3} & \frac{36}{\lambda_0^4} \\ -\frac{2}{\lambda_0^2} & \frac{12}{\lambda_0^3} & \frac{4(2\lambda_0-1)}{\lambda_0^2(1-\lambda_0)^2} & \frac{12(3\lambda_0^2-3\lambda_0+1)}{\lambda_0^3(\lambda_0-1)^3} \\ -\frac{12}{\lambda_0^3} & \frac{36}{\lambda_0^4} & \frac{12(3\lambda_0^2-3\lambda_0+1)}{\lambda_0^3(\lambda_0-1)^3} & \frac{36(4\lambda_0^3-6\lambda_0^2+4\lambda_0-1)}{\lambda_0^4(1-\lambda_0)^4} \end{bmatrix}.$$

Also define $Z^*(v)$ as follows: $Z^*(0) = 0$, $Z^*(v) = Z_1(v)$ for $v < 0$ and $Z^*(v) = Z_2(v)$ for $v > 0$, with

$$\begin{aligned}
Z_1(v) &= (\beta_b)^2|v|^3/3 + v^2\kappa(d)\beta_b\xi_4 + v\sigma^2[2\xi_2'\Omega_1\xi_1 - \xi_1'\Omega_2\xi_1], \quad v < 0, \\
Z_2(v) &= (\beta_b)^2|v|^3/3 + v^2\kappa(d)\beta_b\xi_3 + v\sigma^2[2\xi_2'\Omega_1\xi_1 - \xi_1'\Omega_2\xi_1], \quad v > 0.
\end{aligned}$$

Then, $T^{1/2-d}(\hat{\lambda}-\lambda) \Rightarrow \arg\min_v Z^*(v)$.

4.4 The limiting distribution of other parameters

We turn to the limiting distribution of the other parameters in the models, that is $(\hat{\mu}_1, \hat{\beta}_1, \hat{\beta}_b)$ for Model I, and $(\hat{\mu}_1, \hat{\mu}_b, \hat{\beta}_1, \hat{\beta}_b)$ for Model II.

Theorem 5 *Under assumption A1-A3, the following results hold for all $d \in (-0.5, 0.5)$. 1) For Model I:*

$$\begin{bmatrix} T^{1/2-d}(\hat{\mu}_1 - \mu_1^0) \\ T^{3/2-d}(\hat{\beta}_1 - \beta_1^0) \\ T^{3/2-d}(\hat{\beta}_b - \beta_b^0) \end{bmatrix} \Rightarrow \Sigma_a^{-1}\Sigma_0 \quad \textit{if } m = 0,$$

$$\begin{bmatrix} T^{-1/2-d}(\hat{\mu}_1 - \mu_1^0) \\ T^{1/2-d}(\hat{\beta}_1 - \beta_1^0) \\ T^{1/2-d}(\hat{\beta}_b - \beta_b^0) \end{bmatrix} \Rightarrow \Sigma_a^{-1}\Sigma_1 \quad \textit{if } m = 1,$$

where

$$\Sigma_a^{-1} = \begin{bmatrix} (\lambda_0+3)/\lambda_0 & -3(\lambda_0+1)/\lambda_0^2 & 3/\lambda_0^2(1-\lambda_0) \\ -3(\lambda_0+1)/\lambda_0^2 & 3(3\lambda_0+1)/\lambda_0^3 & -3(2\lambda_0+1)/\lambda_0^3(1-\lambda_0) \\ 3/\lambda_0^2(1-\lambda_0) & -3(2\lambda_0+1)/\lambda_0^3(1-\lambda_0) & 3/\lambda_0^3(1-\lambda_0)^3 \end{bmatrix},$$

$$\Sigma_0 = \kappa(d)\left(\int_0^{\lambda_0} \begin{bmatrix} \frac{3\lambda_0^2-2\lambda_0+6\lambda_0 r-6r}{\lambda_0^2} \\ \frac{\lambda_0^3-\lambda_0+3\lambda_0^2 r-3r}{\lambda_0^2} \\ \frac{-(1-\lambda_0)^2(\lambda_0+3r)}{\lambda_0^2} \end{bmatrix} dW_d(r) + \int_{\lambda_0}^1 \begin{bmatrix} \frac{-3(\lambda_0+1-2r)}{1-\lambda_0} \\ \frac{-\lambda_0^2-3\lambda_0-2+2\lambda_0 r+4r}{1-\lambda_0} \\ -2\lambda_0+4r-2 \end{bmatrix} dW_d(r)\right),$$

and

$$\Sigma_1 = \kappa(d)\left(\int_0^{\lambda_0} \begin{bmatrix} \frac{3(1-\lambda_0)r^2+(3\lambda_0-2)\lambda_0 r-\lambda_0^2}{\lambda_0^2} \\ \frac{(3-2\lambda_0^2)r^2-2\lambda_0(1-2\lambda_0^2)r-\lambda_0^2}{2\lambda_0^2} \\ \frac{(1-\lambda_0)^2(3r^2-2\lambda_0 r-\lambda_0^2)}{2\lambda_0^2} \end{bmatrix} dW_d(r) + \int_{\lambda_0}^1 \begin{bmatrix} \frac{3\{r^2-(1+\lambda_0)r+\lambda_0\}}{1-\lambda_0} \\ \frac{(\lambda_0+2)\{r^2-(1+\lambda_0)r+\lambda_0\}}{1-\lambda_0} \\ 2\{r^2-(1+\lambda_0)r+\lambda_0\} \end{bmatrix} dW_d(r)\right).$$

2) For Model II:

$$\begin{bmatrix} T^{1/2-d}(\hat{\mu}_1 - \mu_1^0) \\ T^{3/2-d}(\hat{\beta}_1 - \beta_1^0) \\ T^{1/2-d}(\hat{\mu}_b - \mu_b^0) - T^{1/2-d}\beta_b(\hat{T}_1 - T_1^0) \\ T^{3/2-d}(\hat{\beta}_b - \beta_b^0) \end{bmatrix} \Rightarrow \kappa(d)\Omega_1 \begin{bmatrix} \int_0^1 dW_d(r) \\ \int_0^1 r dW_d(r) \\ \int_{\lambda_0}^1 dW_d(r) \\ \int_{\lambda_0}^1 (r-\lambda_0) dW_d(r) \end{bmatrix} \quad \textit{if } m = 0,$$

$$\begin{bmatrix} T^{-1/2-d}(\hat{\mu}_1 - \mu_1^0) \\ T^{1/2-d}(\hat{\beta}_1 - \beta_1^0) \\ T^{-1/2-d}(\hat{\mu}_b - \mu_b^0) - T^{-1/2-d}\beta_b(\hat{T}_1 - T_1^0) \\ T^{1/2-d}(\hat{\beta}_b - \beta_b^0) \end{bmatrix} \Rightarrow \kappa(d)\Omega_1 \begin{bmatrix} \int_0^1 W_d(r)dr \\ \int_0^1 rW_d(r)dr \\ \int_{\lambda_0}^1 W_d(r)dr \\ \int_{\lambda_0}^1 (r-\lambda_0)W_d(r)dr \end{bmatrix} \quad \text{if } m = 1.$$

This implies that $\hat{\mu}_b$ is asymptotically unidentified.

Note that except for the unidentified intercept shift $\hat{\mu}_b$, the other parameters, $(\hat{\mu}_0, \hat{\beta}_0, \hat{\beta}_b)$, have the same stochastic order for Model I and II. As noted in PZ, the exact model specification does not matter if one wants to make asymptotic inference on these parameters.

5 Spurious Break

In this section, we consider the properties of the least square estimate of a structural break date when no structural break is present in the data generating process. Nunes et al. (1995) and Bai (1998) showed that the least square estimator of the break date can lead to a spurious break date when the errors are integrated, in the sense that the estimate will not converge to either end of the sample. Kuan and Hsu (1998) considered a change in mean model for a fractionally integrated process with $d^* \in (-0.5, 0.5)$ and showed that a spurious break can be estimated if $d^* \in (0, 0.5)$. Hsu and Kuan (2008) confirmed the possibility of estimating a spurious mean break if the series is a non-stationary fractionally integrated process, i.e., $d^* \in (0.5, 1.5)$. Here, we consider the issue of spurious breaks in the context of Model I with a joint-segmented trend. The DGP is specified as follows; for $t = 1, \ldots, T$,

$$y_t = \mu + \beta t + u_t, \tag{11}$$

and

$$(1-L)^{d^*} u_t \mathbf{1}_{t>0} = \epsilon_t \tag{12}$$

where $\epsilon_t \sim i.i.d.N(0, \sigma^2)$. When estimating a one-break model in slope using Model I, the regression for a candidate break date is

$$y_t = \mu + \delta t + \gamma B_t + u_t, \quad t = 1, \ldots, T.$$

The so-called spurious break problem has been analyzed in the segmented regression model (see, e.g., Nunes et al, 1995). However, we take a global least squares approach. Hence, the theoretical derivations are different. In matrix form, the DGP is

$$Y = X_0\beta + U,$$

and Model I can be written as:

$$Y = X_{T_1}\Gamma + U.$$

The OLS estimate of Γ is $\hat{\Gamma} = (X'_{T_1}X_{T_1})^{-1}X'_{T_1}Y$, the OLS residuals are

$$\hat{U} = [I - X_{T_1}(X'_{T_1}X_{T_1})^{-1}X'_{T_1}]Y,$$

and the sum of squared residuals is given by

$$\hat{U}'\hat{U} = Y'[I - X_{T_1}(X'_{T_1}X_{T_1})^{-1}X'_{T_1}]Y.$$

It is straightforward to show that

$$\begin{aligned}\hat{T}_1 &= \arg\min_{T_1}\ \hat{U}'\hat{U} = \arg\min_{T_1}\ \{\hat{U}'\hat{U} - U'U\} \\ &= \arg\max_{T_1}\ U'X_{T_1}(X'_{T_1}X_{T_1})^{-1}X'_{T_1}U = \arg\max_{T_1}\ M_T(T_1).\end{aligned}$$

Let $M_T^*(T_1)$ be the normalized version of $M_T(T_1)$, that is,

$$\begin{aligned}M_T^*(T_1) &\equiv T^{-2(d+m)}M_T(T_1) \\ &= T^{-(d+m)}U'X_{T_1}D_T^{-1/2}(D_T^{-1/2}X'_{T_1}X_{T_1}D_T^{-1/2})^{-1}D_T^{-1/2}X'_{T_1}UT^{-(d+m)}\end{aligned}$$

where $D_T = diag\{T, T^3, T^3\}$ and $m \in \{0, 1\}$. In order to derive the asymptotic distribution of $M_T^*(T_1)$, we need the following conditions which are similar to those of Nunes et al. (1995) and Bai (1998).

- **Condition S1** There exists a diagonal matrix D_T such that $D_T^{-1/2}X'_{T_1}X_{T_1}D_T^{-1/2} \xrightarrow{p} Q(\lambda)$, uniformly in $\lambda \in (0, 1)$ where $Q(\lambda)$ is assumed to be a positive definite matrix for all $\lambda > 0$.

- **Condition S2** For some $\alpha \geq 0$, $T^{-\alpha/2}D_T^{-1/2}X'_{T_1}U \xrightarrow{p} G(\lambda)$, where $G(\lambda)$ is a stochastic process having continuous sample paths.

Conditions S1 and S2 hold with Assumptions A1 and A2. Note that $\hat{T}_1 = \arg\max_{T_1} M_T(T_1) = \arg\max_{T_1} M_T^*(T_1)$ because the normalization factor $T^{-2(d+m)}$ does not depend on T_1. If Conditions S1 and S2 hold, we have

$$M_T^*(T_1) \xrightarrow{p} M^*(\lambda) \equiv G(\lambda)'Q(\lambda)G(\lambda).$$

It can be shown that (see, Bai, 1997) $\hat{\lambda} \xrightarrow{d} \arg\max_{\lambda\in(0,1)} M^*(\lambda)$. Hence, the estimate of the break fraction $\hat{\lambda}$ is a random variable with support in $(0, 1)$. Note that this is true for all

$d^* \in (-0.5, 0.5) \cup (0.5, 1.5)$, which generalizes the results for $I(0)$ and $I(1)$ processes in Nunes et al. (1995) and Bai (1998).

Below, we show that $M^*(\lambda) \to \infty$ as $\lambda \to \{0, 1\}$ if $d^* \in (-0.5, 0.5)$. Taqqu (1977) showed that for $d^* \in (-0.5, 0.5)$, the fractional Brownian motion $W_d(t),\ t \in \mathbb{R}$ satisfies the following law of iterated logarithms:

$$\limsup_{t \to \infty} \frac{W_d(t)}{(ct^{1+2d} \log\log t)^{1/2}} = 1 \quad a.s.$$

for some positive constant c. Since $W_d(t)$ is self-similar with self-similarity parameter $0.5+d$, for any $c > 0$ it satisfies, $W_d(t) \stackrel{d}{=} c^{-(0.5+d)} W_d(ct)$, where $\stackrel{d}{=}$ denotes equality in distribution. Applying the law of iterated logarithms to $W_d(1/t)$ and self-similarity, we have

$$\limsup_{t \to 0} \frac{W_d(t)}{(ct^{1+2d} \log\log(1/t))^{1/2}} = 1 \quad a.s.$$

Then, for $d \in (-0.5, 0]$,

$$\limsup_{\lambda \to 0} \frac{W_d(\lambda)}{\sqrt{\lambda}} = \limsup_{\lambda \to 0} \sqrt{c\lambda^{2d} \log\log(1/\lambda)} = \infty \quad a.s.$$

It is easy to verify that $W_d(1) - W_d(\lambda)$ is also a fractional Brownian motion $W_d(s)$ with $s = 1 - \lambda$. For $d \in (-0.5, 0]$,

$$\limsup_{\lambda \to 1} \frac{W_d(1) - W_d(\lambda)}{\sqrt{1-\lambda}} = \limsup_{s \to 0} \frac{W_d(s)}{\sqrt{s}} = \infty \quad a.s.$$

Of interest is the behavior of $M^*(\lambda)$ when λ gets closer to either 0 or 1.

Theorem 6 *Under Assumption A1 and A2: 1) for any $d^* \in (-0.5, 0]$, $\limsup_{\lambda \to 0} M^*(\lambda) = \limsup_{\lambda \to 1} M^*(\lambda) = \infty \quad a.s.$; 2) for any $d^* \in (0, 0.5)$, $\limsup_{\lambda \to 1} M^*(\lambda) = \infty \quad a.s.$.*

Theorem 6 implies that no spurious break is estimated if the order of fractional integration is a value in $(-0.5, 0.5)$.

Proposition 1 *For α defined in Condition S2, assume $\alpha \geq 2d^*$ with $d^* \in (0.5, 1.5)$. Then, $\sup_{\lambda \in (0,1)} M^*(\lambda) = O_p(1)$.*

The proof is similar to that of Theorem 1 in Bai (1998) with Lemma 3 and is omitted. Proposition 1 implies that $M^*(\lambda)$ is stochastically bounded for $d^* \in (0.5, 1.5)$. In the following theorem, we show that it is not possible that $\hat{\lambda} \to \{0, 1\}$ in the limit. Note that $M^*(0)$

and $M^*(1)$ are defined without the dummy variable for a slope change B_t in the model. After some algebra, we have

$$M^*(0) = M^*(1) = \kappa^2(d)[4\left(\int_0^1 W_d(r)dr\right)^2 - 12\left(\int_0^1 W_d(r)dr\right)\left(\int_0^1 rW_d(r)dr\right) + 12\left(\int_0^1 rW_d(r)dr\right)^2].$$

Theorem 7 *Under Assumption A1 and A2, for any $d^* \in (0.5, 1.5)$, $M^*(0) = M^*(1) < M^*(\lambda)$, for every $\lambda \in (0, 1)$.*

Theorem 7 implies that the maximum value of $M^*(\lambda)$ cannot be located at 0 or 1 and the value that maximizes $M^*(\lambda)$ on any subset of $[0, 1]$ is bounded away from 0 or 1 since $M^*(\cdot)$ is not a constant process. Hemce, the spurious break feature applies when $d^* \in (0.5, 1.5)$.

5.1 Monte Carlo Experiments

We consider simulation experiments to illustrate the issue of a potential spurious break. The data generating process is specified by

$$y_t = \mu + \beta t + u_t,$$

$$(1 - L)^{d^*} u_t \mathbf{1}_{t>0} = \epsilon_t, \quad \epsilon_t \sim i.i.d.N(0, \sigma^2).$$

for $t = 1, \ldots, T$. Without loss of generality, we set $\mu = \beta = 0$ and we consider $d^* \in \{-0.2, 0.3, 0.8, 1.3\}$. The sample sizes used are $T = 200$ and $T = 2,000$. For each value of d^*, the results are obtained from 10,000 replications. We consider estimating the date of a structural break using either Model I (joint-segmented trend) or Model II (locally disjoint broken trend).

Figure 1 presents the results pertaining to Model I. Figure 1(a) presents histograms of the estimates $\hat{T}_1$ when $T = 200$. For $d^* \in \{-0.2, 0.3\}$, the estimates are concentrated at the two end points (1 and T) indicating that the estimate of the break date is consistent and no spurious break feature is present, consistent with Theorem 6. For $d^* \in \{0.8, 1.3\}$, the estimates of the break date $\hat{T}_1$ tend to cluster near the middle of the sample, which falsely indicates that there is a break in the sample. Figure 1(b) presents histograms of the estimates $\hat{T}_1$ with $T = 2000$. With this larger sample, the estimates often occur near the boundaries, though there is no mass at or very near 0 or 1 with $d^* \in \{0.8, 1.3\}$. Hence, the theoretical results are supported by the simulations.

The corresponding results for Model II are presented in Figure 2. Interestingly, in this case a spurious break occurs for all positive values of d^* even when $T = 2000$. Hence, it appears that simply introducing an irrelevant level shift can alter the results towards having a spurious break. More work is needed to theoretically assess this feature.

These results reinforce the feature discussed in the literature to the effect that structural change and long memory imply similar features in the data, and it is difficult to distinguish one from the other at least in small samples. This suggests the importance of implementing a proper testing procedure for a structural break which should be robust to any a priori unknown order of integration. Recently, Harvey et al. (2009) and Perron and Yabu (2009) suggested testing procedures for a structural change in trend function designed to be robust to $I(0)$ or $I(1)$ errors. Iacone et al. (2013) presented a sup-Wald type test for a change in the slope of a trend function which is robust across fractional values of the order of integration. These tests are useful to avoid the spurious break problem.

6 Conclusion

This paper considered the consistency, rate of convergence and limit distribution of the estimate of a break date in the slope of a linear trend function, with or without a concurrent level shift, when the errors are fractionally integrated with $d^* \in (-0.5, 0.5) \cup (0.5, 1.5)$. Our theoretical results uncover some interesting features. For example, when a concurrent level shift is allowed, the rate of convergence of the estimate of the break date is the same for all values of d^* in the interval $(-0.5, 0.5)$. This feature is linked to the contamination induced by allowed a level shift, previously discussed by Perron and Zhu (2005). In all other cases, the rate of convergence is monotonically decreasing as d^* increases. We also provide results about the so-called spurious break issue and show that it cannot occur in the limit when d^* in the interval $(-0.5, 0.5)$. Simulation experiments illustrate this theoretical result.

The results in this paper can be useful for subsequent work. For instance, Lobato and Velasco (2007) considered efficient Wald test of unit root against a fractionally integrated process with unknown order. However, their procedure does not allow a break under both the null and alternative hypotheses. Accordingly, an interesting avenue would be to extend the Kim and Perron (2009) unit root testing procedure that allows a structural change in the trend function under both the null and alternative hypotheses. Just as the results of Perron and Zhu (2005) and Perron and Yabu (2009) were useful to achieve this task, one could use our results and those of Iacone et al. (2013) to extend the test of Lobato and Velasco (2007). This is currently the object of ongoing research.

Appendix

We consider the proofs of Theorems 1-4 for Models I and II separately, for ease of exposition. Note first that all limit statements are taken as $T \to \infty$.

A.1 Results for Model I

Model I can be represented in matrix notation as

$$Y = X_{T_1}\gamma + U$$
$$= \begin{bmatrix} \iota & \mathbf{t} & \mathbf{B}_{T_1} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \beta_1 \\ \beta_b \end{bmatrix} + U$$

where $Y = (y_1, \ldots, y_T)'$, $U = (u_1, \ldots, u_T)'$, $X_{T_1} = (x(T_1)_1, \ldots, x(T_1)_T)'$ with $x(T_1)_t' = (1, t, B_t)$ and $\gamma = (\mu_1, \beta_1, \beta_b)'$. Note that the matrix X_{T_1} depends on the candidate value of the break date T_1. In the proof, we only consider the case $T_1 > T_1^0$. It is straightforward to apply the same arguments to the case where $T_1 < T_1^0$. For $T_1 > T_1^0$, let

$$\widetilde{\iota}_b(t) = \begin{cases} 0 & \text{if } 1 \leq t \leq T_1^0 \\ \frac{t - T_1^0}{T_1 - T_1^0} & \text{if } T_1^0 < t < T_1 \\ 1 & \text{if } T_1 \leq t \leq T. \end{cases}$$

and for $T_1 = T_1^0$, let

$$\widetilde{\iota}_b(t) = \iota_b(t) = \begin{cases} 0 & \text{if } 1 \leq t \leq T_1^0 \\ 1 & \text{if } T_1^0 < t \leq T. \end{cases}$$

With this notation, we can write

$$(X_{T_1^0} - X_{T_1})\gamma = \beta_b(T_1 - T_1^0)\widetilde{\iota}_b$$

Note that $\widetilde{\iota}_b([Tr])$ converges to a continuous function $f_{\widetilde{\iota}_b}(r)$ over $[0, 1]$ defined by, for $\lambda > \lambda_0$,

$$f_{\widetilde{\iota}_b}(r) = \begin{cases} 0 & \text{if } 0 \leq r \leq \lambda_0 \\ \frac{r - \lambda_0}{\lambda - \lambda_0} & \text{if } \lambda_0 < r < \lambda \\ 1 & \text{if } \lambda \leq r \leq 1, \end{cases}$$

and by, for $\lambda = \lambda_0$,

$$f_{\widetilde{\iota}_b}(r) = f_{\iota_b}(r) = \begin{cases} 0 & \text{if } 0 \leq r \leq \lambda_0 \\ 1 & \text{if } \lambda_0 < r \leq 1. \end{cases}$$

Pertaining to the proof of Theorem 1, we first consider the term (XX). We have

$$\begin{aligned}(XX) &= \gamma_0'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})(X_{T_1^0} - X_{T_1})\gamma_0 \\ &= (T_1 - T_1^0)^2\beta_b^2\widetilde{\iota}_b'(I - P_{T_1})\widetilde{\iota}_b\end{aligned}$$

where the second equality holds because the first two columns of $(X_{T_1^0} - X_{T_1})$ are zeros by construction. Note that $\widetilde{\iota}_b'(I - P_{T_1})\widetilde{\iota}_b$ is the sum of squared residuals from a regression $\widetilde{\iota}_b$ on $[\iota \quad \mathbf{t} \quad \mathbf{B}_{T_1}]$. Define

$$S_T = \widetilde{\iota}_b'(I - P_{T_1})\widetilde{\iota}_b.$$

Next, consider the continuous time least-squares projection of the function $f_{\widetilde{\iota}_b}(r)$ on $[1 \quad r \quad f_B(r)]$, where $f_B(r) = (r-\lambda)\mathbf{1}_{r\geq\lambda}$. Let $[\hat{\alpha} \quad \hat{\beta} \quad \hat{\psi}]$ denote the estimates of the coefficients and let S_∞ denote the resulting SSR. From the definition of a Riemann integral, $T^{-1}S_T \to S_\infty$, where

$$S_\infty = \int_0^1 \left(f_{\widetilde{\iota}_b}(r) - \hat{\alpha} - \hat{\beta}r - \hat{\psi}f_B(r)\right)^2 dr.$$

Suppose that $\hat{\alpha} = \hat{\beta} = 0$. It is easy to show that $S_\infty > 0$ from the definition of $f_{\widetilde{\iota}_b}(r)$ and $f_B(r)$. Otherwise, we have

$$S_\infty \geq \int_0^{\min\{\lambda,\lambda_0\}} \left(f_{\widetilde{\iota}_b}(r) - \hat{\alpha} - \hat{\beta}r - \hat{\psi}f_B(r)\right)^2 dr = \int_0^{\min\{\lambda,\lambda_0\}} (\hat{\alpha} + \hat{\beta}r)^2 dr > 0$$

where the equality holds because of the definition on $f_{\widetilde{\iota}_b}(r)$ and $f_B(r)$ and the fact that both λ and λ_0 are bounded away from zero. Hence, $0 < S_\infty < \infty$ and $S_T = O(T)$. Accordingly,

$$(XX) = (T_1 - T_1^0)^2\beta_b^2 O(T).$$

Next, we consider the term (XU). We have

$$\begin{aligned}(XU) &= \gamma'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})U \\ &= \beta_b(T_1 - T_1^0)\widetilde{\iota}_b'(I - P_{T_1})U\end{aligned}$$

Define $\widetilde{f}_{\widetilde{\iota}_b}(r)$ as the projection residuals from a least-squares regression of $f_{\widetilde{\iota}_b}(r)$ on $[1 \quad r \quad f_B(r)]$. By the properties of orthogonal projections and the result for (XX), we have

$$\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)dr = \int_0^1 \left(f_{\widetilde{\iota}_b}(r) - \hat{\alpha} - \hat{\beta}r - \hat{\psi}f_B(r)\right) dr = 0$$

and

$$\int_0^1 [\widetilde{f}_{\widetilde{\iota}_b}(r)]^2 dr = S_\infty = O(1)$$

uniformly over all λ. By the functional central limit theorem (FCLT) and the continuous mapping theorem,

$$T^{-(d+1/2)}\widetilde{\iota}_b'(I - P_k)U \Rightarrow \kappa(d)\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)dW_d(r) \quad \text{if } m = 0.$$

Similarly,

$$T^{-(d+3/2)}\widetilde{\iota}_b'(I-P_k)U \Rightarrow \kappa(d)\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)W_d(r)dr \quad \text{if } m=1,$$

where $\kappa^2(d) = [b_\psi^2\Gamma(1-2d)E\varepsilon_0^2]/(1+2d)\Gamma(1+d)\Gamma(1-d)$. We deduce that

$$E[\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)dW_d(r)] = 0 \quad \text{and} \quad E[\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)W_d(r)dr] = 0,$$

$$\begin{aligned} Var[\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)dW_d(r)] &= \int_0^1\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)\widetilde{f}_{\widetilde{\iota}_b}(u)E[dW_d(u)dW_d(r)] \\ &= \alpha_d\int_0^1\int_0^1 |u-r|^{2d-1}\widetilde{f}_{\widetilde{\iota}_b}(r)\widetilde{f}_{\widetilde{\iota}_b}(u)dudr = \|\widetilde{f}_{\widetilde{\iota}_b}\|_{\mathcal{H}}^2, \quad \text{if } m=0 \end{aligned}$$

where $\alpha_d = 2d(d+1/2)$ and $\mathcal{H}$ is a Banach space, and

$$\begin{aligned} Var[\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)W_d(r)dr] &= \int_0^1\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)\widetilde{f}_{\widetilde{\iota}_b}(u)E[W_d(r)W_d(u)]dudr \\ &= \int_0^1\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)\widetilde{f}_{\widetilde{\iota}_b}(u)\frac{1}{2}(|u|^{2d+1}+|r|^{2d+1}-|u-r|^{2d+1})dudr \\ &= \int_0^1\int_0^r \widetilde{f}_{\widetilde{\iota}_b}(r)\widetilde{f}_{\widetilde{\iota}_b}(u)(|u|^{2d+1}+|r|^{2d+1}-|u-r|^{2d+1})dudr \\ &= O_p(1), \quad \text{if } m=1 \end{aligned}$$

uniformly over all $\lambda \in [\pi, 1-\pi]$. Therefore, $\int_0^1 \widetilde{f}_{\widetilde{\iota}_b}(r)W_d(r)dr = O_p(1)$ and

$$\widetilde{\iota}_b'(I-P_{T_1})U = \begin{cases} O_p(T^{d+1/2}) & \text{if } m=0, \\ O_p(T^{d+3/2}) & \text{if } m=1. \end{cases}$$

Hence, we have

$$(XU) = \begin{cases} \beta_b(T_1-T_1^0)O_p(T^{d+1/2}) & \text{if } m=0, \\ \beta_b(T_1-T_1^0)O_p(T^{d+3/2}) & \text{if } m=1. \end{cases}$$

Finally, we consider the term (UU). Define $D_T = diag(T^{d+1/2}, T^{d+3/2}, T^{d+3/2})$ with $d \in (-0.5, 0.5)$. We have

$$\begin{aligned} (UU) &= U'(P_{T_1^0}-P_{T_1})U \\ &= U\{X_{T_1^0}(X_{T_1^0}'X_{T_1^0})^{-1}X_{T_1^0}' - X_{T_1}(X_{T_1}'X_{T_1})^{-1}X_{T_1}'\}U \\ &= U(X_{T_1^0}-X_{T_1})D_T^{-1}[D_T^{-1}X_{T_1^0}'X_{T_1^0}D_T^{-1}]^{-1}D_T^{-1}X_{T_1^0}'U \\ &\quad + UX_{T_1}D_T^{-1}[D_T^{-1}X_{T_1}'X_{T_1}D_T^{-1}]^{-1}D_T^{-1}[X_{T_1}'X_{T_1}-X_{T_1^0}'X_{T_1^0}] \\ &\times D_T^{-1}[D_T^{-1}X_{T_1^0}'X_{T_1^0}D_T^{-1}]^{-1}D_T^{-1}X_{T_1^0}'U \\ &\quad + UX_{T_1}D_T^{-1}[D_T^{-1}X_{T_1}'X_{T_1}D_T^{-1}]^{-1}D_T^{-1}(X_{T_1^0}-X_{T_1})'U. \end{aligned}$$

Applying the FCLT for $d \in (-0.5, 0.5)$ and $m = 0$,

$$T^{-(d+1/2)} \sum_{t=1}^{T} u_t \Rightarrow \kappa(d) W_d(1),$$

$$T^{-(d+3/2)} \sum_{t=1}^{T} t u_t \Rightarrow \kappa(d)[W_d(1) - \int_0^1 W_d(r) dr] = \kappa(d) \int_0^1 r dW_d(r).$$

Also, from Lemma 3 with $m = 1$.

$$T^{-(d+3/2)} \sum_{t=1}^{T} u_t \Rightarrow \kappa(d) \int_0^1 W_d(r) dr,$$

$$T^{-(d+5/2)} \sum_{t=1}^{T} t u_t \Rightarrow \kappa(d) \int_0^1 r W_d(r) dr.$$

In addition, it is easy to show that

$$T^{-3} \sum_{t=T_1+1}^{T} (t - T_1)^2 \rightarrow \int_\lambda^1 (r - \lambda)^2 dr,$$

$$T^{-3} \sum_{t=T_1+1}^{T} (t - T_1) t \rightarrow \int_\lambda^1 (r - \lambda) r dr,$$

$$T^{-2} \sum_{t=T_1+1}^{T} (t - T_1) \rightarrow \int_\lambda^1 (r - \lambda) dr.$$

We next consider each term in (UU).

1. $D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1}$ and $D_T^{-1} X_{T_1^0}' X_{T_1^0} D_T^{-1}$ are $O(T^{-2d})$ uniformly in λ.

2. When $m = 0$, $D_T^{-1} X_{T_1}' U$ and $D_T^{-1} X_{T_1^0}' U$ are $O_p(1)$ uniformly in λ, and

$$D_T^{-1} X_{T_1}' U = \begin{bmatrix} T^{-(d+1/2)} \sum_{t=1}^{T} u_t \\ T^{-(d+3/2)} \sum_{t=1}^{T} t u_t \\ T^{-(d+3/2)} \sum_{t=T_1+1}^{T} (t - T_1) u_t \end{bmatrix} \Rightarrow \begin{bmatrix} \kappa(d) W_d(1) \\ \kappa(d) \int_0^1 r dW_d(r) \\ \kappa(d) \int_\lambda^1 (r - \lambda) dW_d(r) \end{bmatrix}.$$

   When $m = 1$, $D_T^{-1} X_{T_1}' U$ and $D_T^{-1} X_{T_1^0}' U$ are $O_p(T)$ uniformly in λ, and

$$T^{-1} D_T^{-1} X_{T_1}' U = \begin{bmatrix} T^{-(d+3/2)} \sum_{t=1}^{T} u_t \\ T^{-(d+5/2)} \sum_{t=1}^{T} t u_t \\ T^{-(d+5/2)} \sum_{t=T_1+1}^{T} (t - T_1) u_t \end{bmatrix} \Rightarrow \begin{bmatrix} \kappa(d) \int_0^1 W_d(r) dr \\ \kappa(d) \int_0^1 r W_d(r) dr \\ \kappa(d) \int_\lambda^1 (r - \lambda) W_d(r) dr \end{bmatrix}.$$

3. $U'(X_{T_1^0} - X_{T_1})D_T^{-1}$. It suffices to consider the third column of $(X_{T_1^0} - X_{T_1})$ because the first two columns are zeros. We have

$$\begin{aligned} T^{-(d+1/2+m)}U'(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1}) &= T^{-(d+1/2+m)} \sum_{T_1^0+1}^{T_1} (t - T_1^0)u_t + T^{-(d+1/2+m)}(T_1 - T_1^0) \sum_{T_1+1}^{T} u_t \\ &= |T_1 - T_1^0| O_p(1) \quad \text{for } m \in \{0, 1\}. \end{aligned}$$

4. $D_T^{-1}[X'_{T_1}X_{T_1} - X'_{T_1^0}X_{T_1^0}]D_T^{-1}$. As noted earlier, it suffices to consider the terms in which $\mathbf{B}_{T_1}$ and $\mathbf{B}_{T_1^0}$ are involved.

$$\begin{aligned} \mathbf{B}'_{T_1^0}\mathbf{B}_{T_1^0} - \mathbf{B}'_{T_1}\mathbf{B}_{T_1} &= |T_1 - T_1^0| O(T^2) \\ \mathbf{B}'_{T_1^0}\mathbf{t} - \mathbf{B}'_{T_1}\mathbf{t} &= |T_1 - T_1^0| O(T^2) \\ \mathbf{B}'_{T_1^0}\iota - \mathbf{B}'_{T_1}\iota &= |T_1 - T_1^0| O(T) \end{aligned}$$

Hence, we have

$$D_T^{-1}[X'_{T_1}X_{T_1} - X'_{T_1^0}X_{T_1^0}]D_T^{-1} = |T_1 - T_1^0| O(T^{-(1+2d)}), \quad \text{for } m \in \{0, 1\}.$$

Based on the results 1-4,

$$(UU) = \begin{cases} |T_1 - T_1^0| O_p(T^{-1+2d}) & \text{if } m = 0 \\ |T_1 - T_1^0| O_p(T^{1+2d}) & \text{if } m = 1 \end{cases}$$

This completes the proof of Theorem 1 for Model I.

A.1.1 Proof of Consistency (Theorem 2)

From the proof of Theorem 1, we know that for Model I, if $m = 0$,

$$\begin{aligned} (\hat{X}\hat{X}) &= (T_1^0 - \hat{T}_1)^2 \beta_b^2 O(T) \\ (\hat{X}\hat{U}) &= \beta_b (T_1^0 - \hat{T}_1) O_p(T^{1/2+d}) \\ (\hat{U}\hat{U}) &= |T_1^0 - \hat{T}_1| O_p(T^{-1+2d}). \end{aligned}$$

and, if $m = 1$,

$$\begin{aligned} (\hat{X}\hat{X}) &= (T_1^0 - \hat{T}_1)^2 \beta_b^2 O(T) \\ (\hat{X}\hat{U}) &= \beta_b (T_1^0 - \hat{T}_1) O_p(T^{3/2+d}) \\ (\hat{U}\hat{U}) &= |T_1^0 - \hat{T}_1| O_p(T^{1+2d}) \end{aligned}$$

We consider the proof for $m = 0$ (the proof for $m = 1$ is similar). Suppose that $\hat{\lambda} \nrightarrow_p \lambda_0$. Then, the results above imply that $(\hat{X}\hat{X}) = O(T^3)$, $(\hat{X}\hat{U}) = O_p(T^{3/2+d})$, and $(\hat{U}\hat{U}) = O_p(T^{2d})$ for $d \in (-0.5, 0.5)$. Therefore, for sufficiently large T, the term $(\hat{X}\hat{X})$ dominates the others with some probability. It implies that the key inequality $(\hat{X}\hat{X}) + 2(\hat{X}\hat{U}) + (\hat{U}\hat{U}) \leq 0$ cannot hold with probability 1. Since this inequality is valid for all T, we have a contradiction. Hence, we can conclude that $\hat{\lambda} \rightarrow_p \lambda_0$.

A.1.2 Rate of Convergence (Theorem 3)

Consider the set

$$V(\epsilon) = \{T_1 : |T_1 - T_1^0| < \epsilon T, \forall \epsilon > 0\}.$$

From the consistency of $\hat{T}_1$ in Theorem 2, $\Pr\left(\hat{T}_1 \in V(\epsilon)\right) \to 1$ as $T \to \infty$. Hence, it suffices to consider the behavior of $S(T_1)$ for all $T_1 \in V(\epsilon)$. Consider another set $V_c(\epsilon)$ defined by

$$\begin{aligned} V_c(\epsilon) &= \{T_1 : |T_1 - T_1^0| < \epsilon T \quad \text{and} \quad |T_1 - T_1^0| > CT^{-1/2+d+m}, \\ \forall \epsilon &> 0, \forall d \in (-0.5, 0.5), m = \{0, 1\}\}. \end{aligned}$$

Note that $V_c(\epsilon) \subset V(\epsilon)$. Since $S(\hat{T}_1) \leq S(T_1^0)$ with probability 1, we can claim that $T_1^0 \notin V_c(\epsilon)$ by showing that for each $\eta > 0$, there exists a constant $C > 0$ such that

$$\Pr\left(\min_{T_1 \in V_c(\epsilon)} \{S(T_1) - S(T_1^0)\} \leq 0\right) < \eta \tag{A.1}$$

Equation (A.1) implies that a minimum cannot be obtained in the set $V_c(\epsilon)$ and that $|T_1 - T_1^0| \leq CT^{-1/2+d+m}$ must hold with an arbitrary large probability. Equation (A.1) is equivalent to

$$\Pr\left(\min_{T_1 \in V_c(\epsilon)} \{(XX) + 2(XU) + (UU)\} \leq 0\right) < \eta$$

Based on results derived in Theorem 1, we can apply the following normalizations. If $m = 0$, then

$$\begin{aligned} \frac{(XX)}{|T_1 - T_1^0| T^{1/2+d}} &= \frac{|T_1 - T_1^0|^2 \beta_b^2 O(T)}{|T_1 - T_1^0| T^{1/2+d}} > \frac{CT^{-1/2+d} \beta_b^2 O(T)}{T^{-1/2+d} T} = aC + o(1) \\ \frac{(XU)}{|T_1 - T_1^0| T^{1/2+d}} &= \frac{|T_1 - T_1^0| \beta_b O_p(T^{1/2+d})}{|T_1 - T_1^0| T^{1/2+d}} = O_p(1) \\ \frac{(UU)}{|T_1 - T_1^0| T^{1/2+d}} &= \frac{|T_1 - T_1^0| O_p(T^{-1+2d})}{|T_1 - T_1^0| T^{1/2+d}} = o_p(1). \end{aligned}$$

If $m = 1$, then

$$\begin{aligned} \frac{(XX)}{|T_1 - T_1^0| T^{3/2+d}} &= \frac{|T_1 - T_1^0|^2 \beta_b^2 O(T)}{|T_1 - T_1^0| T^{3/2+d}} > \frac{CT^{1/2+d} \beta_b^2 O(T)}{T^{1/2+d} T} = aC + o(1) \\ \frac{(XU)}{|T_1 - T_1^0| T^{3/2+d}} &= \frac{|T_1 - T_1^0| \beta_b O_p(T^{3/2+d})}{|T_1 - T_1^0| T^{3/2+d}} = O_p(1) \\ \frac{(UU)}{|T_1 - T_1^0| T^{3/2+d}} &= \frac{|T_1 - T_1^0| O_p(T^{1+2d})}{|T_1 - T_1^0| T^{3/2+d}} = o_p(1). \end{aligned}$$

where a is a positive constant. Here, we simply use the fact that $|T_1 - T_1^0| < \epsilon T$ and $|T_1 - T_1^0| > CT^{-1/2+d+m}$ in $V_c(\epsilon)$. Therefore, Equation (A.1) is satisfied for all $\epsilon > 0$ if we choose a sufficiently large $C > 0$.

A.1.3 Limiting Distribution of the estimate of the break date

Consider first the case with $m = 1$. Define the set $D(C) = \{T_1 : |T_1 - T_1^0| < CT^{1/2+d}\}$, for some positive number C, and $m_T = T^{-1/2-d}|T_1 - T_1^0|$. We are interested in the stochastic orders of the other parameters. We analyze

$$\arg\min_{T_1 \in D(C)} [SSR(T_1) - SSR(T_1^0)].$$

For $T_1 \in D(C)$, we have $|T_1 - T_1^0| = O(T^{1/2+d})$. Hence, $(XX) = |T_1 - T_1^0|^2 O(T) = O(T^{2+2d})$, $(XU) = |T_1 - T_1^0| O_p(T^{3/2+d}) = O_p(T^{2+2d})$ and $(UU) = |T_1 - T_1^0| O_p(T^{1+2d}) = O_p(T^{3/2+3d})$. Then,

$$\begin{aligned}\arg\min_{T_1 \in D(C)} [SSR(T_1) - SSR(T_1^0)] &= \arg\min_{T_1 \in D(C)} [(XX) + 2(XU) + (UU)]/T^{2+2d} \\ &= \arg\min_{T_1 \in D(C)} [(XX)/T^{2+2d} + 2(XU)/T^{2+2d} + o_p(1)],\end{aligned}$$

hence we only need to consider the first two terms. Note that on the set $D(C)$, $|\lambda - \lambda_0| = O(T^{-1/2+d})$ for $d \in (-0.5, 0.5)$. Using this fact, we can derive the following results that will subsequently be applied:

$$\begin{aligned} T^{2d} D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1} &= \begin{bmatrix} 1 & 1/2 & (1-\lambda_0)^2/2 \\ 1/2 & 1/3 & (1-\lambda_0)^2(2+\lambda_0)/6 \\ (1-\lambda_0)^2/2 & (1-\lambda_0)^2(2+\lambda_0)/6 & (1-\lambda_0)^3/3 \end{bmatrix} + o(1) \\ &\equiv \Sigma_a + o(1), \end{aligned}$$

and the inverse is $T^{-2d}(D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} = \Sigma_a^{-1} + o(1)$ with

$$\Sigma_a^{-1} = \begin{bmatrix} (\lambda_0+3)/\lambda_0 & -3(\lambda_0+1)/\lambda_0^2 & 3/\lambda_0^2(1-\lambda_0) \\ -3(\lambda_0+1)/\lambda_0^2 & 3(3\lambda_0+1)/\lambda_0^3 & -3(2\lambda_0+1)/\lambda_0^3(1-\lambda_0) \\ 3/\lambda_0^2(1-\lambda_0) & -3(2\lambda_0+1)/\lambda_0^3(1-\lambda_0) & 3/\lambda_0^3(1-\lambda_0)^3 \end{bmatrix}$$

We have

$$\begin{aligned} (XX) &= \beta_b^2 (\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'(I - P_{T_1})(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1}) \\ &= \beta_b^2 \{(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1}) \\ &\quad -(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})' X_{T_1} D_T^{-1} (D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} D_T^{-1} X_{T_1}' (\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})\} \end{aligned}$$

The second term in (XX) is such that

$$\begin{aligned} T^{-1}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})' X_{T_1} D_T^{-1} &= |T_1 - T_1^0| T^{-1/2-d} T^{-1/2+d} \tilde{\iota}_b' X_{T_1} D_T^{-1} \\ &= m_T \begin{bmatrix} 1-\lambda_0 & \frac{1-\lambda_0^2}{2} & \frac{(1-\lambda_0)^2}{2} \end{bmatrix} + o(1) \end{aligned}$$

where $m_T = T^{-1/2-d}|T_1 - T_1^0|$. Using the results above,

$$T^{-1-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})' X_{T_1} D_T^{-1} (D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} = m_T \begin{bmatrix} -\frac{1-\lambda_0}{2} & \frac{3(1-\lambda_0)}{2\lambda_0} & \frac{3(2\lambda_0-1)}{2\lambda_0(1-\lambda_0)} \end{bmatrix} + o(1) \tag{A.2}$$

Hence,

$$T^{-2-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})' X_{T_1} (X_{T_1}' X_{T_1})^{-1} X_{T_1}' (\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1}) = \left[\frac{(1-\lambda_0)\lambda_0}{4}\right] m_T^2 + o(1) \tag{A.3}$$

and

$$\begin{aligned} T^{-2-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1}) &= T^{-2-2d}|T_1 - T_1^0|^2 \widetilde{\iota}_b' \widetilde{\iota}_b \\ &= m_T^2 T^{-1} \widetilde{\iota}_b' \widetilde{\iota}_b \\ &= (1-\lambda_0) m_T^2 + o(1) \end{aligned}$$

Combining (A.2) and (A.3), we obtain

$$T^{-2-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'(I - P_{T_1})(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1}) = \left[\frac{(1-\lambda_0)\lambda_0}{4}\right] m_T^2 + o(1)$$

Now,

$$(XU) = \gamma_0 (X_{T_1^0} - X_{T_1})'(I - P_{T_1})U = \delta(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'(I - P_{T_1})U.$$

We have,

$$\begin{aligned} T^{-2-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'U &= |T_1 - T_1^0| T^{-1/2-d} T^{-3/2-d} \widetilde{\iota}_b' U \\ &= m_T \kappa(d) \int_\lambda^1 (1 - r/\lambda) W_d(r) dr + o_p(1), \end{aligned}$$

$$\begin{aligned} & T^{-2-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})' X_{T_1} (X_{T_1}' X_{T_1})^{-1} X_{T_1}' U \\ &= T^{-2-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})' X_{T_1} D_T^{-1} (D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} D_T^{-1} X_{T_1}' U \\ &= T^{-1}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})' X_{T_1} D_T^{-1} T^{-2d} (D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} T^{-1} D_T^{-1} X_{T_1}' U, \end{aligned}$$

and

$$\begin{aligned} T^{-1} D_T^{-1} X_{T_1}'^{-1} U &= T^{-1} \left[T^{-1/2-d} \sum_{t=1}^T u_t \quad T^{-3/2-d} \sum_{t=1}^T t u_t \quad T^{-3/2-d} \sum_{t=T_1+1}^T (t - T_1) u_t \right]' \\ &= \left[T^{-3/2-d} \sum_{t=1}^T u_t \quad T^{-5/2-d} \sum_{t=1}^T t u_t \quad T^{-5/2-d} \sum_{t=T_1^0+1}^T (t - T_1^0) u_t + o_p(1) \right]'. \end{aligned}$$

Hence, for $d \in (-0.5, 0.5)$ and $m = 1$, we have

$$\begin{aligned}
&T^{-2-2d}(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'(I - P_{T_1})U \\
&= \kappa(d)\{\int_{\lambda_0}^1 W_d(r)dr + \frac{1-\lambda_0}{2}\int_0^1 W_d(r)dr - \\
&\quad - \frac{3(1-\lambda_0)}{2\lambda_0}\int_0^1 rW_d(r)dr - \frac{3(2\lambda_0-1)}{2\lambda_0(1-\lambda_0)}\int_{\lambda_0}^1 (r-\lambda_0)W_d(r)dr\}\beta_b m_T + o_p(1) \\
&= \kappa(d)\beta_b m_T \int_{\lambda_0}^1 W_d^*(r)dr + o_p(1),
\end{aligned}$$

where $W_d^*(r)$ is the residuals function from a continuous time least-squares regression of $W_d(r)$ on $\{1, r, (r-\lambda_0)\mathbf{1}_{r>\lambda_0}\}$. Therefore,

$$\begin{aligned}
m_T^* &= \arg\min{}_{m_T \in D(C)}[(XX)T^{-2-2d} + 2(XU)T^{-2-2d} + o_p(1)] \\
&= \arg\min{}_{m_T \in D(C)}[m_T^2\beta_b^2\frac{\lambda_0(1-\lambda_0)}{4} + 2\kappa(d)m_T\beta_b\int_{\lambda_0}^1 W_d^*(r)dr] + o_p(1)
\end{aligned}$$

by the continuous mapping theorem. Note that the objective function does not change if $T_1 - T_1^0 < 0$. We can conclude that

$$m_T^* = T^{-1/2-d}|\hat{T}_1 - T_1^0| \Rightarrow -\frac{4\kappa(d)\int_{\lambda_0}^1 W_d^*(r)dr}{\lambda_0(1-\lambda_0)\beta_b}.$$

Next, consider the case with $m = 0$. Define $m_T = T^{1/2+d}|T_1 - T_1^0|$ for this case. Note that $T^{-1/2-d}\widetilde{\iota}_b' U \Rightarrow \kappa(d)\int_{\lambda_0}^1 dW_d(r)$. For (XX), we have the same results as for $m = 1$. For (XU), we have:

$$\begin{aligned}
(XU) &= \beta_b(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1})'(I - P_{T_1})U \\
&= T^{-1/2-d}\beta_b m_T \widetilde{\iota}_b'(I - P_{T_1})U \\
&= T^{-1/2-d}\beta_b m_T \widetilde{\iota}_b'^{-1/2-d}\beta_b m_T \widetilde{\iota}_b' X_{T_1} D_T^{-1}(D_T^{-1}X_{T_1}'X_{T_1}D_T^{-1})^{-1}D_T^{-1}X_{T_1}U
\end{aligned}$$

$$\begin{aligned}
&= \beta_b m_T \kappa(d)\left[\int_{\lambda_0}^1 dW_d(r)\right. \\
&\quad \left. - \begin{bmatrix} \frac{\lambda_0-1}{2} & \frac{3(1-\lambda_0)}{2\lambda_0} & \frac{3(2\lambda_0-1)}{2\lambda_0(1-\lambda_0)} \end{bmatrix} \begin{bmatrix} \int_0^1 dW_d(r) \\ \int_0^1 r dW_d(r) \\ \int_{\lambda_0}^1 (r-\lambda_0)dW_d(r) \end{bmatrix}\right] + o_p(1) \\
&= \beta_b m_T \kappa(d)\left[\int_0^{\lambda_0} \frac{\lambda_0 - \lambda_0^2 + 3r - 3r\lambda_0}{2\lambda_0}dW_d(r) + \int_{\lambda_0}^1 \frac{\lambda_0(2+\lambda_0-3r)}{2(1-\lambda_0)}dW_d(r)\right] + o_p(1) \\
&\equiv \beta_b m_T \kappa(d)\zeta + o_p(1).
\end{aligned}$$

For (UU), we know that U is an $I(d)$ process with $d \in (-0.5, 0.5)$. It is easy to show that $U'(X_{T_1^0} - X_{T_1})D_T^{-1} = |T_1 - T_1^0|O_p(T^{-1})$, $D_T^{-1}X'_{T_1^0}U = O_p(1)$, and $D_T^{-1}X'_{T_1^0}X_{T_1^0}D_T^{-1} = O_p(T^{-2d})$. Hence, $(UU) = |T_1 - T_1^0|O_p(T^{-1-2d})$ which is dominated by (XU) asymptotically. The optimal m_T^* is therefore given by

$$m_T^* = T^{3/2+d}(\hat{\lambda} - \lambda_0) \Rightarrow \frac{-4\kappa(d)\zeta}{\beta_b \lambda_0 (1 - \lambda_0)}.$$

A.1.4 Limit Distributions of the Other Parameters

The OLS estimates of the regression coefficients γ is

$$\begin{aligned}
\hat{\gamma} &= (X'_{\hat{T}_1}X_{\hat{T}_1})^{-1}X'_{\hat{T}_1}Y \\
&= (X'_{\hat{T}_1}X_{\hat{T}_1})^{-1}X'_{\hat{T}_1}X_{T_1^0}\gamma_0 + (X'_{\hat{T}_1}X_{\hat{T}_1})^{-1}X'_{\hat{T}_1}U \\
&= \gamma_0 + D_T^{-1}(D_T^{-1}X'_{\hat{T}_1}X_{\hat{T}_1}D_T^{-1})^{-1}D_T^{-1}X'_{\hat{T}_1}(X_{T_1^0} - X_{\hat{T}_1})\gamma_0 \\
&+ D_T^{-1}(D_T^{-1}X'_{\hat{T}_1}X_{\hat{T}_1}D_T^{-1})^{-1}D_T^{-1}X'_{\hat{T}_1}U.
\end{aligned}$$

Hence,

$$D_T(\hat{\gamma} - \gamma_0) = (D_T^{-1}X'_{\hat{T}_1}X_{\hat{T}_1}D_T^{-1})^{-1}[D_T^{-1}X'_{\hat{T}_1}(X_{T_1^0} - X_{\hat{T}_1})\gamma_0 + D_T^{-1}X'_{\hat{T}_1}U].$$

First, for $m = 0$,

$$\begin{aligned}
& D_T^{-1}X'_{\hat{T}_1}(X_{T_1^0} - X_{\hat{T}_1})\gamma_0 + D_T^{-1}X'_{\hat{T}_1}U \\
= \;& D_T^{-1}X'_{\hat{T}_1}\beta_b|T_1 - T_1^0|\widetilde{\iota}_b + D_T^{-1}X'_{\hat{T}_1}U \\
= \;& D_T^{-1}X'_{\hat{T}_1}\beta_b|T_1 - T_1^0|T^{1/2-d}\widetilde{\iota}_b T^{-1/2+d} + D_T^{-1}X'_{\hat{T}_1}U \\
= \;& m_T\beta_b T^{-1/2+d}D_T^{-1}X'_{\hat{T}_1}\widetilde{\iota}_b + D_T^{-1}X'_{\hat{T}_1}U
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow \frac{-4\kappa(d)\zeta}{\beta_b\lambda_0(1-\lambda_0)}\beta_b \begin{bmatrix} 1-\lambda_0 \\ \frac{1-\lambda_0^2}{2} \\ \frac{(1-\lambda_0)^2}{2} \end{bmatrix} + \kappa(d) \begin{bmatrix} \int_0^1 dW_d(r) \\ \int_0^1 r dW_d(r) \\ \int_{\lambda_0}^1 (r-\lambda_0) dW_d(r) \end{bmatrix} \\
&= \kappa(d) \left(\int_0^{\lambda_0} \begin{bmatrix} \frac{3\lambda_0^2 - 2\lambda_0 + 6\lambda_0 r - 6r}{\lambda_0^2} \\ \frac{\lambda_0^3 - \lambda_0 + 3\lambda_0^2 r - 3r}{\lambda_0^2} \\ \frac{-(1-\lambda_0)^2(\lambda_0 + 3r)}{\lambda_0^2} \end{bmatrix} dW_d(r) + \int_{\lambda_0}^1 \begin{bmatrix} \frac{-3(\lambda_0 + 1 - 2r)}{1-\lambda_0} \\ \frac{-\lambda_0^2 - 3\lambda_0 - 2 + 2\lambda_0 r + 4r}{1-\lambda_0} \\ -2\lambda_0 + 4r - 2 \end{bmatrix} dW_d(r) \right) \\
&\equiv \Sigma_0.
\end{aligned}$$

Since $T^{-2d}(D_T^{-1}X'_{T_1}X_{T_1}D_T^{-1})^{-1} \xrightarrow{p} \Sigma_a^{-1}$,

$$T^{-2d}D_T(\hat{\gamma} - \gamma_0) \Rightarrow \Sigma_a^{-1}\Sigma_0.$$

Second, for $m = 1$,

$$T^{-1}D_T(\hat{\gamma} - \gamma_0) = (D_T^{-1}X'_{\hat{T}_1}X_{\hat{T}_1}D_T^{-1})^{-1}[T^{-1}D_T^{-1}X'_{\hat{T}_1}(X_{T_1^0} - X_{\hat{T}_1})\gamma_0 + T^{-1}D_T^{-1}X'_{\hat{T}_1}U].$$

Then, we have

$$\begin{aligned}
&T^{-1}D_T^{-1}X'_{\hat{T}_1}(X_{T_1^0} - X_{\hat{T}_1})\gamma_0 + T^{-1}D_T^{-1}X'^{-1}_{\hat{T}_1}D_T^{-1}X'_{\hat{T}_1}\beta_b|T_1 - T_1^0|\widetilde{\iota}_b + T^{-1}D_T^{-1}X'_{\hat{T}_1}U \\
&= D_T^{-1}X'_{\hat{T}_1}\beta_b|T_1 - T_1^0|T^{-1/2-d}\widetilde{\iota}_b T^{-1/2+d} + T^{-1}D_T^{-1}X'_{\hat{T}_1}U \\
&= m_T\beta_b T^{-1/2+d}D_T^{-1}X'_{\hat{T}_1}\widetilde{\iota}_b + T^{-1}D_T^{-1}X'_{\hat{T}_1}U \\
&\Rightarrow -\kappa(d)\int_{\lambda_0}^1 W_d^*(d)dr \begin{bmatrix} \frac{4}{\lambda_0} \\ \frac{2(1+\lambda_0)}{\lambda_0} \\ \frac{2(1-\lambda_0)}{\lambda_0} \end{bmatrix} + \kappa(d)\begin{bmatrix} \int_0^1 W_d(r)dr \\ \int_0^1 rW_d(r)dr \\ \int_{\lambda_0}^1 (r-\lambda)W_d dr(r) \end{bmatrix} \\
&= \kappa(d)\left(\int_0^{\lambda_0} \begin{bmatrix} \frac{3(1-\lambda_0)r^2 + (3\lambda_0-2)\lambda_0 r - \lambda_0^2}{\lambda_0^2} \\ \frac{(3-2\lambda_0^2)r^2 - 2\lambda_0(1-2\lambda_0^2)r - \lambda_0^2}{2\lambda_0^2} \\ \frac{(1-\lambda_0)^2(3r^2 - 2\lambda_0 r - \lambda_0^2)}{2\lambda_0^2} \end{bmatrix} dW_d(r) + \int_{\lambda_0}^1 \begin{bmatrix} \frac{3\{r^2-(1+\lambda_0)r+\lambda_0\}}{1-\lambda_0} \\ \frac{(\lambda_0+2)\{r^2-(1+\lambda_0)r+\lambda_0\}}{1-\lambda_0} \\ 2\{r^2 - (1+\lambda_0)r + \lambda_0\} \end{bmatrix} dW_d(r) \right) \\
&\equiv \Sigma_1.
\end{aligned}$$

Therefore,

$$T^{-1-2d}D_T(\hat{\gamma} - \gamma_0) \Rightarrow \Sigma_a^{-1}\Sigma_1.$$

A.2 Results for Model II

We now consider results for Model II. The proofs of the consistency is similar to that for Model I. In any event, the relevant bound will be derived in the proof of the limit distribution.

A.2.1 Consistency (Theorem 2)

From Theorem 1, for $m = 0$:

$$\begin{aligned}
(\hat{X}\hat{X}) &= (T_1^0 - \hat{T}_1)^3\beta_b^2 O(1) \\
(\hat{X}\hat{U}) &= \beta_b(T_1^0 - \hat{T}_1)^{3/2+d}O_p(1) \\
(\hat{U}\hat{U}) &= |T_1^0 - \hat{T}_1|^{1/2+d}O_p(T^{-1/2+d}),
\end{aligned}$$

and for $m = 1$:

$$\begin{aligned}
(\hat{X}\hat{X}) &= (T_1^0 - \hat{T}_1)^3\beta_b^2 O(1) \\
(\hat{X}\hat{U}) &= \beta_b(T_1^0 - \hat{T}_1)^2 O_p(T^{1/2+d}) \\
(\hat{U}\hat{U}) &= |T_1^0 - \hat{T}_1|O_p(T^{1+2d}).
\end{aligned}$$

The proof of consistency is similar to that for Model I. Suppose that $\hat{\lambda} \overset{p}{\nrightarrow} \lambda$. Then, with $m = 1$, $(\hat{X}\hat{X}) = O(T^3)$, $(\hat{X}\hat{U}) = O_p(T^{5/2+d})$ and $(\hat{U}\hat{U}) = O_p(T^{2+2d})$ for all $d \in (-0.5, 0.5)$. Hence, with some positive probability, $(\hat{X}\hat{X})$ dominate the other two terms, so that this result cannot be compatible with the key inequality (10). Hence, we have a contradiction and conclude that $\hat{\lambda} \overset{p}{\rightarrow} \lambda$.

A.2.2 Rate of Convergence (Theorem 3)

We consider the set

$$\begin{aligned} \widetilde{V}_c(\epsilon) &= \{T_1 : |T_1 - T_1^0| < \epsilon T \quad \text{and} \quad |T_1 - T_1^0| > CT^{m(d+1/2)}, \\ \forall \epsilon &> 0, \forall d \in (-0.5, 0.5), m = \{0, 1\}\}. \end{aligned}$$

Given the results in Theorem 1, if $m = 0$:

$$\begin{aligned} \frac{(XX)}{|T_1 - T_1^0|^{3/2+d}} &= \frac{|T_1 - T_1^0|^3 \beta_b^2 O(1)}{|T_1 - T_1^0|^{3/2+d}} = |T_1 - T_1^0|^{3/2-d} \beta_b^2 O(1) > C^{3/2-d} \beta_b^2 O(1) = aC_1 + o(1), \\ \frac{(XU)}{|T_1 - T_1^0|^{3/2+d}} &= \frac{|T_1 - T_1^0|^{3/2+d} \beta_b O_p(1)}{|T_1 - T_1^0|^{3/2+d}} = O_p(1), \\ \frac{(UU)}{|T_1 - T_1^0|^{3/2+d}} &= \frac{|T_1 - T_1^0|^{1/2+d} O_p(T^{-1/2+d})}{|T_1 - T_1^0|^{3/2+d}} = o_p(1), \end{aligned}$$

and if $m = 1$:

$$\begin{aligned} \frac{(XX)}{|T_1 - T_1^0|^2 T^{1/2+d}} &= \frac{|T_1 - T_1^0|^3 \beta_b^2 O(1)}{|T_1 - T_1^0|^2 T^{1/2+d}} > \frac{CT^{1/2+d} \beta_b^2 O(1)}{T^{1/2+d}} = aC + o(1), \\ \frac{(XU)}{|T_1 - T_1^0|^2 T^{1/2+d}} &= \frac{|T_1 - T_1^0|^2 \beta_b O_p(T^{1/2+d})}{|T_1 - T_1^0|^2 T^{1/2+d}} = O_p(1), \\ \frac{(UU)}{|T_1 - T_1^0|^2 T^{1/2+d}} &= \frac{|T_1 - T_1^0| O_p(T^{1+2d})}{|T_1 - T_1^0|^2 T^{1/2+d}} = o_p(1) \end{aligned}$$

for $\forall d \in (-0.5, 0.5)$ where a is a positive constant. It is easy to show that

$$Pr\left(\min_{T_1 \in \widetilde{V}_c(\epsilon)} \{S(T_1) - S(T_1^0)\} \leq 0 \right) < \eta$$

for any $\epsilon > 0$ if we choose a sufficiently large $C > 0$. This completes the proof.

A.2.3 Limit Distribution of the Estimate of the Break Date

Given the results in Theorem 3, we work on the set $D_0(C) = \{T_1 : |T_1 - T_1^0| < T^d C\}$ if $m = 0$ and $D_1(C) = \{T_1 : |T_1 - T_1^0| < T^{1/2+d} C\}$ if $m = 1$, for some positive C. In other words, for $\lambda = T_1/T$, $|\lambda - \lambda_0| = O_p(T^{-1+d})$ with $m = 0$ and $|\lambda - \lambda_0| = O_p(T^{-1/2+d})$ with $m = 1$. In matrix notation, Model II can be expressed as

$$Y = X_{T_1} \gamma + U$$

with

$$X_{T_1} = \begin{bmatrix} \iota & \mathbf{t} & \mathbf{C}_{T_1} & \mathbf{B}_{T_1} \end{bmatrix}$$

where $\iota = (1,\ldots,1)'$, $\mathbf{t} = (1,2,\ldots,T)'$, $\mathbf{C}_{T_1} = (C_1,\ldots,C_T)$, $\mathbf{B}_{T_1} = (B_1,\ldots,B_T)'$ and $\gamma = [\mu_1 \quad \beta_1 \quad \mu_b \quad \beta_b]'$. For $T_1 > T_1^0$,

$$\mathbf{C}_{T_1^0} - \mathbf{C}_{T_1} = \begin{cases} 1 & \text{if } T_1^0 \le t \le T_1 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1} - (T_1 - T_1^0)\mathbf{C}_{T_1} = \begin{cases} t - T_1^0 & \text{if } T_1^0 \le t \le T_1 \\ 0 & \text{otherwise.} \end{cases}$$

When $T_1 < T_1^0$,

$$\mathbf{C}_{T_1^0} - \mathbf{C}_{T_1} = \begin{cases} -1 & \text{if } T_1^0 \le t \le T_1 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1} - (T_1 - T_1^0)\mathbf{C}_{T_1} = \begin{cases} -(t - T_1^0) & \text{if } T_1^0 \le t \le T_1 \\ 0 & \text{otherwise.} \end{cases}$$

We shall use the following notations. For $T_1^0 > T_1$,

$$g_1(T_1 - T_1^0) = \sum_{t=T_1+1}^{T_1^0} [\mu_b + \beta_b(t - T_1^0)],$$

$$h_1(T_1 - T_1^0) = \sum_{t=T_1+1}^{T_1^0} [\mu_b + \beta_b(t - T_1^0)]^2$$

and for $T_1^0 < T_1$,

$$g_2(T_1 - T_1^0) = \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)],$$

$$h_2(T_1 - T_1^0) = \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)]^2$$

We first consider the term (XX). Noting that $(T_1 - T_1^0)(I - P_{T_1})C_{T_1} = 0$, we have

$$(XX) = \gamma_0'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})(X_{T_1^0} - X_{T_1})\gamma_0$$
$$= [(\mathbf{C}_{T_1^0} - \mathbf{C}_{T_1})\mu_b + (\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1} - (T_1 - T_1^0)\mathbf{C}_{T_1})\beta_b)]'(I - P_{T_1})$$
$$\times [(\mathbf{C}_{T_1^0} - \mathbf{C}_{T_1})\mu_b + (\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1} - (T_1 - T_1^0)\mathbf{C}_{T_1})\beta_b)]$$
$$= \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)]^2$$
$$- \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)]x(T_1)_t' D_T^{-1}(D_T^{-1}X_{T_1}'X_{T_1}D_T^{-1})^{-1}D_T^{-1} \sum_{t=T_1^0+1}^{T_1} x(T_1)_t[\mu_b + \beta_b(t - T_1^0)]$$

where $D_T = diag(T^{1/2+d}, T^{3/2+d}, T^{1/2+d}, T^{3/2+d})$. Note that for $T_1 > T_1^0$,

$$\sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)]x(T_1)_t' D_T^{-1}$$
$$= T^{-d} \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)][T^{-1/2} \quad tT^{-3/2} \quad 0 \quad 0]$$
$$= T^{-1/2-d} \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)][1 \quad t/T \quad 0 \quad 0]$$
$$= T^{-1/2-d} \sum_{k=1}^{T_1-T_1^0} [\mu_b + \beta_b k][1 \quad (k + T_1^0)/T \quad 0 \quad 0]$$
$$= T^{-1/2-d} g_2[1 \quad T_1^0/T \quad 0 \quad 0] + T^{-1/2-d} \sum_{k=1}^{T_1-T_1^0} [\mu_b + \beta_b k][0 \quad k/T \quad 0 \quad 0]$$
$$\le T^{-1/2-d}|g_2|[1 \quad T_1^0/T \quad 0 \quad 0] + |g_2|T^{-1/2-d}\frac{|T_1 - T_1^0|}{T}[0 \quad 1 \quad 0 \quad 0]$$
$$= O_p(|g_2|T^{-1/2-d}).$$

where the last step follows from the fact that $|T_1 - T_1^0|/T \xrightarrow{p} 0$ on both $D_0(C)$ and $D_1(C)$. Also,

$$(D_T^{-1}X_{T_1}'X_{T_1}D_T^{-1})^{-1} = O_p(T^{2d}).$$

Hence, the second term in (XX) is such that

$$\gamma_0'(X_{T_1^0} - X_{T_1})'P_{T_1}(X_{T_1^0} - X_{T_1})\gamma_0 = O_p(g_2^2 T^{-1}) = o_p(h_2)$$

because $|\lambda - \lambda_0| = O_p(T^{-1+d})$ if $m = 0$ and $|\lambda - \lambda_0| = O_p(T^{-1/2+d})$ if $m = 1$ where $d \in (-0.5, 0.5)$. Therefore,

$$(XX) = \begin{cases} h_2 + o_p(h_2) & \text{if } T_1 > T_1^0 \\ h_1 + o_p(h_1) & \text{if } T_1 \leq T_1^0 \end{cases}$$

This implies that

$$(XX) = |T_1 - T_1^0|^3 O(1)$$

since μ_b is fixed. Consider now the term (XU). For $m = 1$, we have

$$\begin{aligned} (XU) &= \gamma'(X_{T_1^0} - X_{T_1})'(I - P_{T_1})U \\ &= \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)]u_t \\ &\quad - [\sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)]x(T_1)_t' D_T^{-1}](D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} D_T^{-1} X_{T_1}' U \end{aligned}$$

We consider each term of this expression.

1. When $T_1^0 < T_1$, let $u_t = u_{T_1^0} + v_k$. Then,

$$\begin{aligned} T^{-1/2-d} \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)]u_t &= T^{-1/2-d} \sum_{k=1}^{T_1 - T_1^0} [\mu_b + \beta_b k] u_{T_1^0+k} \\ &= T^{-1/2-d} g_2 u_{T_1^0} + T^{-3/2-d} \sum_{k=1}^{T_1-T_1^0} [\mu_b + \beta_b k] v_k \\ &= g_2 \kappa(d) W_d(\lambda_0) + o_p(g_2). \end{aligned}$$

2. $T^{2d} D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1} = \Omega_1^{-1} + o(1)$, where

$$\Omega_1 = \begin{bmatrix} \frac{4}{\lambda_0} & -\frac{6}{\lambda_0^2} & \frac{2}{\lambda_0} & \frac{6}{\lambda_0^2} \\ -\frac{6}{\lambda_0^2} & \frac{12}{\lambda_0^3} & -\frac{6}{\lambda_0^2} & -\frac{12}{\lambda_0^3} \\ \frac{2}{\lambda_0} & -\frac{6}{\lambda_0^2} & \frac{4}{\lambda_0(1-\lambda_0)} & \frac{6(1-2\lambda_0)}{\lambda_0^2(1-\lambda_0)^2} \\ \frac{6}{\lambda_0^2} & -\frac{12}{\lambda_0^3} & \frac{6(1-2\lambda_0)}{\lambda_0^2(1-\lambda_0)^2} & \frac{12(3\lambda_0^2-3\lambda_0+1)}{\lambda_0^3(1-\lambda_0)^3} \end{bmatrix},$$

and

$$\Omega_1^{-1} = \begin{bmatrix} 1 & \frac{1}{2} & 1-\lambda_0 & \frac{(1-\lambda_0)^2}{2} \\ \frac{1}{2} & \frac{1}{3} & \frac{(1-\lambda_0)^2}{2} & \frac{(1-\lambda_0)^2(2+\lambda_0)}{6} \\ 1-\lambda_0 & \frac{1-\lambda_0^2}{2} & 1-\lambda_0 & \frac{(1-\lambda_0)^2}{2} \\ \frac{(1-\lambda_0)^2}{2} & \frac{(1-\lambda_0)^2(2+\lambda_0)}{6} & \frac{(1-\lambda_0)^2}{2} & \frac{(1-\lambda_0)^3}{3} \end{bmatrix}.$$

3. $T^{-1}D_T^{-1}X'_{T_1}U \Rightarrow \kappa(d)\xi_1$, where

$$\xi_1 = \begin{bmatrix} \int_0^1 W_d(r)dr \\ \int_0^1 rW_d(r)dr \\ \int_0^{\lambda_0} W_d(r)dr \\ \int_{\lambda_0}^1 (r-\lambda_0)W_d(r)dr \end{bmatrix} = \begin{bmatrix} \int_0^1 (1-r)dW_d(r) \\ \int_0^1 \frac{1-r^2}{2}dW_d(r) \\ \int_0^{\lambda_0}(1-\lambda_0)dW_d(r) + \int_{\lambda_0}^1 (1-r)dW_d(r) \\ \int_0^{\lambda_0} \frac{(1-\lambda_0)^2}{2}dW_d(r) + \int_{\lambda_0}^1 \frac{(1-\lambda_0)^2-(1-\lambda_0)^2}{2}dW_d(r) \end{bmatrix}$$

using integration by parts.

4. When $T_1^0 < T_1$,

$$\begin{aligned} & T^{-1/2-d} \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t-T_1^0)]X_{T_1}D_T^{-1} \\ &= T^{-1-2d} \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t-T_1^0)][1 \quad t/T \quad 0 \quad 0] \\ &= T^{-1-2d}g_2[1 \quad \lambda_0 \quad 0 \quad 0] + T^{-1-2d} \sum_{k=1}^{T_1-T_1^0} [\mu_b + \beta_b k][0 \quad k/T \quad 0 \quad 0] \\ &= T^{-1-2d}g_2[1 \quad \lambda_0 \quad 0 \quad 0] + o_p(g_2T^{-1-2d}). \end{aligned}$$

Combining the results 1-4, we obtain that

$$\begin{aligned} (XU) &= T^{1/2+d}\{g_2\kappa(d)W_d(\lambda_0) - g_2[1 \quad \lambda_0 \quad 0 \quad 0]\Omega_1\kappa(d)\xi_1 + o_p(1)\} \\ &= T^{1/2+d}g_2\kappa(d)\{W_d(\lambda_0) - [1 \quad \lambda_0 \quad 0 \quad 0]\Omega_1\xi_1 + o_p(1)\} \\ &= T^{1/2+d}g_2\kappa(d)\xi_3 + o_p(T^{1/2+d}g_2). \end{aligned}$$

After some algebra, we have

$$\xi_3 = W_d(\lambda_0) - [1 \quad \lambda_0 \quad 0 \quad 0]\Omega_1\xi_1 = \int_0^{\lambda_0} \left(\frac{3r^2 - 2\lambda_0 r}{\lambda_0^2}\right) dW_d(r).$$

We can show that when $T_1^0 > T_1$,

$$T^{-1/2-d} \sum_{t=T_1+1}^{T_1^0} [\mu_b + \beta_b(t - T_1^0)]u_t = g_1\kappa(d)W_d(\lambda_0) + o_p(g_1)$$

and

$$T^{-1/2-d} \sum_{t=T_1+1}^{T_1^0} [\mu_b + \beta_b(t - T_1^0)]x(T_1)_t' D_T^{-1} = T^{-1-2d} g_1 [1 \quad \lambda_0 \quad 1 \quad 0] + o_p(g_1 T^{-1-2d}).$$

Hence,

$$(XU) = T^{1/2+d} g_1 \kappa(d) \xi_4 + o_p(T^{1/2+d} g_1)$$

where

$$\xi_4 = \int_{\lambda_0}^1 [(r-1)(3r - 2\lambda_0 - 1)/(1-\lambda_0)^2] dW_d(r).$$

These results imply that

$$(XU) = T^{1/2+d}\kappa(d) \begin{cases} g_2\xi_3 & \text{if } T_1^0 < T_1 \\ g_1\xi_4 & \text{if } T_1^0 > T_1 \end{cases} + o_p(1)$$

and

$$(XU) = |T_1 - T_1^0|^2 O_p(T^{1/2+d}).$$

We finally consider the term (UU). We have

$$\begin{aligned}
&(UU) = U'(P_{T_1^0} - P_{T_1})U \\
&= U'(X_{T_1^0} - X_{T_1}) D_T^{-1} (D_T^{-1} X_{T_1^0}' X_{T_1^0} D_T^{-1})^{-1} D_T^{-1} X_{T_1^0}' U \\
&\quad + U' X_{T_1} D_T^{-1} (D_T^{-1} X_{T_1^0}' X_{T_1^0} D_T^{-1})^{-1} D_T^{-1} [X_{T_1}' X_{T_1} - X_{T_1^0}' X_{T_1^0}] \\
&\times D_T^{-1} (D_T^{-1} X_{T_1^0}' X_{T_1^0} D_T^{-1})^{-1} D_T^{-1} X_{T_1^0}' U \\
&\quad + U' X_{T_1} D_T^{-1} (D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} D_T^{-1} (X_{T_1^0} - X_{T_1})' U
\end{aligned}$$

We first have

$$T^{-1/2-d} U'(\mathbf{C}_{\mathbf{T_1^0}} - \mathbf{C}_{\mathbf{T_1}}) = \kappa(d)(T_1^0 - T_1) \int_{\lambda_0}^1 W_d(r)dr + o_p(1),$$

$$T^{-3/2-d} U'(\mathbf{B}_{\mathbf{T_1^0}} - \mathbf{B}_{\mathbf{T_1}}) = \kappa(d)(T_1^0 - T_1) \int_{\lambda_0}^1 rW_d(r)dr + o_p(1).$$

Hence,

$$U'(X_{T_1^0} - X_{T_1}) D_T^{-1} = (T_1 - T_1^0)[\kappa(d)\xi_2' + o_p(1)]$$

where $\xi_2' = [0 \quad 0 \quad \int_{\lambda_0}^1 W_d(r)dr \quad \int_{\lambda_0}^1 rW_d(r)dr]$. For the second term in (UU), we have

$$D_T^{-1}[X_{T_1}'X_{T_1} - X_{T_1^0}'X_{T_1^0}]D_T^{-1} = -(T_1 - T_1^0)T^{-1-2d}\Sigma_f$$

with

$$\Sigma_f = \begin{bmatrix} 0 & 0 & 1 & 1-\lambda_0 \\ 0 & 0 & \lambda_0 & \frac{1-\lambda_0^2}{2} \\ 1 & \lambda_0 & 1 & 1-\lambda_0 \\ 1-\lambda_0 & \frac{1-\lambda_0^2}{2} & 1-\lambda_0 & (1-\lambda_0)^2 \end{bmatrix}.$$

Hence,

$$\begin{aligned} &T^{1+2d}U'X_{T_1}D_T^{-1}(D_T^{-1}X_{T_1^0}'X_{T_1^0}D_T^{-1})^{-1}D_T^{-1}[X_{T_1}'X_{T_1} - X_{T_1^0}'X_{T_1^0}] \\ &\times D_T^{-1}(D_T^{-1}X_{T_1^0}'X_{T_1^0}D_T^{-1})^{-1}D_T^{-1}X_{T_1^0}'U \\ &= -(T_1 - T_1^0)T^{1+2d}\kappa^2(d)[\xi_1'\Omega_2\xi_1 + o_p(1)] \end{aligned}$$

where

$$\Omega_2 = \Omega_1^{-1}\Sigma_f\Omega_1^{-1} = \begin{bmatrix} -\frac{4}{\lambda_0^2} & \frac{12}{\lambda_0^3} & -\frac{2}{\lambda_0^2} & -\frac{12}{\lambda_0^3} \\ \frac{12}{\lambda_0^3} & -\frac{36}{\lambda_0^4} & \frac{12}{\lambda_0^3} & \frac{36}{\lambda_0^4} \\ -\frac{2}{\lambda_0^2} & \frac{12}{\lambda_0^3} & \frac{4(2\lambda_0-1)}{\lambda_0^2(1-\lambda_0)^2} & \frac{12(3\lambda_0^2-3\lambda_0+1)}{\lambda_0^3(\lambda_0-1)^3} \\ -\frac{12}{\lambda_0^3} & \frac{36}{\lambda_0^4} & \frac{12(3\lambda_0^2-3\lambda_0+1)}{\lambda_0^3(\lambda_0-1)^3} & \frac{36(4\lambda_0^3-6\lambda_0^2+4\lambda_0-1)}{\lambda_0^4(1-\lambda_0)^4} \end{bmatrix}.$$

Collecting the results above, we have

$$(UU) = (T_1 - T_1^0)T^{1+2d}\kappa^2(d)[2\xi_2'\Omega_1\xi_1 - \xi_1'\Omega_2\xi_1 + o_p(1)].$$

This implies that with $m = 1$,

$$(UU) = |T_1 - T_1^0|O_p(T^{1+2d}).$$

Define $m_T = (T_1 - T_1^0)T^{-1/2-d}$. It is easy to show that both h_1 and h_2 are asymptotically equivalent to $T^{3/2+3d}(\beta_b)^2|m_T|^3/3$ and both g_1 and g_2 are asymptotically equivalent to $T^{1+2d}m_T^2\beta_b/2$, therefore

$$\begin{aligned} T^{-3/2-3d}(XX) &= \beta_b^2|m_T|^3/3 + o_p(1), \\ 2T^{-3/2-3d}(XU) &= \begin{cases} \kappa(d)m_T^2\beta_b\xi_3 + o_p(1) & \text{if } m_T > 0 \\ \kappa(d)m_T^2\beta_b\xi_4 + o_p(1) & \text{if } m_T < 0 \end{cases} \\ T^{-3/2-3d}(UU) &= m_T\kappa(d)^2[2\xi_2'\Omega_1\xi_1 - \xi_1'\Omega_2\xi_1] + o_p(1). \end{aligned}$$

Define $Z^*(v;\lambda_0,\beta_b,\kappa(d))$ as follows: $Z^*(0)=0$, $Z^*(v)=Z_1(v)$ for $v<0$ and $Z^*(v)=Z_2(v)$ for $v>0$, with

$$Z_1^*(v;\lambda_0,\beta_b,\kappa(d)) = (\beta_b)^2|v|^3/3 + v^2\kappa(d)\beta_b\xi_4 + v\kappa(d)^2[2\xi_2'\Omega_1\xi_1 - \xi_1'\Omega_2\xi_1] + o_p(1),$$
$$Z_2^*(v;\lambda_0,\beta_b,\kappa(d)) = (\beta_b)^2|v|^3/3 + v^2\kappa(d)\beta_b\xi_3 + v\kappa(d)^2[2\xi_2'\Omega_1\xi_1 - \xi_1'\Omega_2\xi_1] + o_p(1).$$

By the continuous mapping theorem, we have

$$m_T^* \equiv (\hat{T}_1 - T_1^0)T^{-1/2-d} \Rightarrow \arg\min{}_v Z^*(v;\lambda_0,\beta_0,\kappa).$$

Now, consider the case with $d\in(-0.5,0.5)$ and $m=0$. The following argument applies to the set

$$D_0(C) = \{T_1 : |T_1 - T_1^0| < T^dC\}$$

and accordingly we have $|\lambda-\lambda_0| = O_p(T^{-1+d})$ for $\lambda = T_1/T$. As in the case with $m=1$,

$$(XX) = \begin{cases} h_1 + o_p(h_1) & \text{if } T_1 < T_1^0 \\ h_2 + o_p(h_2) & \text{if } T_1 > T_1^0. \end{cases}$$

If $T_1 > T_1^0$,

$$\begin{aligned}(XU) = {} & \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t-T_1^0)]u_t \\ & - \{\sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t-T_1^0)]x(T_1)_t'D_T^{-1}\}(D_T^{-1}X_{T_1}'X_{T_1}D_T^{-1})^{-1}D_T^{-1}X_{T_1}'U.\end{aligned}$$

We next consider each term of (XU).

1.

$$\begin{aligned}\sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t-T_1^0)]u_t &= \sum_{k=1}^{T_1-T_1^0} [\mu_b + \beta_b k]u_{k+T_1^0} \\ &= \sum_{k=1}^{T_1-T_1^0} \mu_b u_{k+T_1^0} + \beta_b \sum_{k=1}^{T_1-T_1^0} k u_{k+T_1^0} = O_p(|T_1 - T_1^0|^{3/2+d}).\end{aligned}$$

2.

$$
\begin{aligned}
&\sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)] x(T_1)_t' D_T^{-1} \\
&= T^{-1/2-d} \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)][1 \quad t/T \quad 0 \quad 0] \\
&= T^{-1/2-d} \sum_{k=1}^{T_1 - T_1^0} [\mu_b + \beta_b k][1 \quad (k + T_1^0)/T \quad 0 \quad 0] \\
&= T^{-1/2-d} \sum_{k=1}^{T_1 - T_1^0} [\mu_b + \beta_b k][1 \quad k/T \quad 0 \quad 0] + T^{-1/2-d} \sum_{k=1}^{T_1 - T_1^0} [\mu_b + \beta_b k][1 \quad \lambda_0 \quad 0 \quad 0] \\
&= O_p(|T_1 - T_1^0|^2 T^{-1/2-d}).
\end{aligned}
$$

3. $(D_T^{-1} X_{T_1}' X_{T_1} D_T^{-1})^{-1} = O_p(T^{2d})$.

4. $D_T^{-1} X_{T_1}' U = O_p(1)$.

Since we search in a set for which $|T_1 - T_1^0| < T^d C$ for some $C > 0$ and $|\lambda - \lambda_0| = O_p(T^{-1+d})$, $\gamma_0'(X_{T_1^0} - X_{\hat{T}_1})' P_{T_1} U$ is dominated by $\gamma_0'(X_{T_1^0} - X_{\hat{T}_1})' U$ asymptotically. Hence,

$$
(XU) = |T_1 - T_1^0|^{3/2+d} O_p(1).
$$

We can derive the results for $T_1^0 > T_1$ in a similar way. In sum,

$$
(XU) = \begin{cases} \sum_{t=T_1^0+1}^{T_1} [\mu_b + \beta_b(t - T_1^0)] u_t + o_p(1) & \text{if } T_1 > T_1^0 \\ 0 & \text{if } T_1 = T_1^0 \\ \sum_{t=T_1+1}^{T_1^0} [\mu_b + \beta_b(t - T_1^0)] u_t + o_p(1) & \text{if } T_1 < T_1^0. \end{cases}
$$

Next, consider the term (UU). We have

$$
\begin{aligned}
T^{-1/2-d} U'(\mathbf{C}_{T_1^0} - \mathbf{C}_{T_1}) &= T^{-1/2-d} \sum_{t=\min\{T_1,T_1^0\}+1}^{\max\{T_1,T_1^0\}} u_t \\
&= T^{-1/2-d} |T_1 - T_1^0|^{1/2+d} |T_1 - T_1^0|^{-1/2-d} \sum_{t=\min\{T_1,T_1^0\}+1}^{\max\{T_1,T_1^0\}} u_t \\
&= T^{-1/2-d} |T_1 - T_1^0|^{1/2+d} O_p(1).
\end{aligned}
$$

and

$$T^{-3/2-d}U'(\mathbf{B}_{T_1^0} - \mathbf{B}_{T_1}) = T^{-1}(T^{-1/2-d}U'\mathbf{B}_{T_1^0} - T^{-1/2-d}U'\mathbf{B}_{T_1})$$
$$= T^{-1}|T_1^0 - T_1|O_p(1).$$

Hence,

$$U'(X_{T_1^0} - X_{T_1})D_T^{-1} = |T_1 - T_1^0|^{1/2+d}O_p(T^{-1/2-d}).$$

Then following the same arguments as for Model I, we have

$$(UU) = |T_1 - T_1^0|^{1/2+d}O_p(T^{-1/2-d})O_p(T^{2d})O_p(1)$$
$$= |T_1 - T_1^0|^{1/2+d}O_p(T^{-1/2+d}).$$

Following Bai (1997), we define a stochastic process $S^*(\nu)$ on the set of integers as follows:

$$S^*(\nu) = \begin{cases} S_1(\nu) & \text{if } \nu < 0 \\ 0 & \text{if } \nu = 0 \\ S_2(\nu) & \text{if } \nu > 0 \end{cases}$$

with

$$S_1(\nu) = \sum_{k=\nu+1}^{0} (\mu_b + \beta_b k)^2 - 2\sum_{k=\nu+1}^{0} (\mu_b + \beta_b k)u_k, \quad \nu = -1, -2, \ldots,$$
$$S_2(\nu) = \sum_{k=1}^{\nu} (\mu_b + \beta_b k)^2 - 2\sum_{k=1}^{\nu} (\mu_b + \beta_b k)u_k, \quad \nu = 1, 2, \ldots.$$

Under the assumption that u_t is strictly stationary and has a continuous distribution, the rest of the proof is similar to that of Bai (1997, p.592) and, hence omitted.

A.2.4 Limit Distributions of the Other Parameters

As for Model I, we use the facts that

$$D_T(\hat{\gamma} - \gamma_0) = (D_T^{-1}X'_{\hat{T}_1}X_{\hat{T}_1}D_T^{-1})^{-1}[D_T^{-1}X'_{\hat{T}_1}(X_{T_1^0} - X_{\hat{T}_1})\gamma_0 + D_T^{-1}X'_{\hat{T}_1}U],$$

and

$$T^{-2d}(D_T^{-1}X'_{\hat{T}_1}X_{\hat{T}_1}D_T^{-1})^{-1} = \Omega_1 + o(1)$$

where

$$\Omega_1 = \begin{bmatrix} \frac{4}{\lambda_0} & -\frac{6}{\lambda_0^2} & \frac{2}{\lambda_0} & \frac{6}{\lambda_0^2} \\ -\frac{6}{\lambda_0^2} & \frac{12}{\lambda_0^3} & -\frac{6}{\lambda_0^2} & -\frac{12}{\lambda_0^3} \\ \frac{2}{\lambda_0} & -\frac{6}{\lambda_0^2} & \frac{4}{\lambda_0(1-\lambda_0)} & \frac{6(1-2\lambda_0)}{\lambda_0^2(1-\lambda_0)^2} \\ \frac{6}{\lambda_0^2} & -\frac{12}{\lambda_0^3} & \frac{6(1-2\lambda_0)}{\lambda_0^2(1-\lambda_0)^2} & \frac{12(3\lambda_0^2-3\lambda_0+1)}{\lambda_0^3(1-\lambda_0)^3} \end{bmatrix}.$$

Hence, we obtain

$$
\begin{aligned}
&T^{-2d}D_T(\hat{\gamma}-\gamma_0)\\
&=\Omega_1^{-1}\left(\beta_b|\hat{T}_1-T_1^0|T^{1/2-d}\begin{bmatrix}1-\lambda_0\\ \frac{1-(\lambda_0)^2}{2}\\ 1-\lambda_0\\ \frac{(1-\lambda_0)^2}{2}\end{bmatrix}+\kappa(d)\begin{bmatrix}\int_0^1 dW_d(r)\\ \int_0^1 rdW_d(r)\\ \int_{\lambda_0}^1 dW_d(r)\\ \int_{\lambda_0}^1(r-\lambda_0)dW_d(r)\end{bmatrix}\right)+o_p(1)\\
&=\beta_b|\hat{T}_1-T_1^0|T^{1/2-d}\begin{bmatrix}0\\0\\1\\0\end{bmatrix}+\kappa(d)\Omega_1^{-1}\begin{bmatrix}\int_0^1 dW_d(r)\\ \int_0^1 rdW_d(r)\\ \int_{\lambda_0}^1 dW_d(r)\\ \int_{\lambda_0}^1(r-\lambda_0)dW_d(r)\end{bmatrix}+o_p(1).
\end{aligned}
$$

Note that the limiting distribution of $\hat{\mu}_b$ depends on that of $|\hat{T}_1-T_1^0|$. Similarly, it is easy to show that, when $m=1$,

$$
\begin{aligned}
&T^{-1-2d}D_T(\hat{\gamma}-\gamma_0)\\
&=\beta_b|\hat{T}_1-T_1^0|T^{-1/2-d}\begin{bmatrix}0\\0\\1\\0\end{bmatrix}+\kappa(d)\Omega^{-1}\begin{bmatrix}\int_0^1 W_d(r)dr\\ \int_0^1 rW_d(r)dr\\ \int_{\lambda_0}^1 W_d(r)dr\\ \int_{\lambda_0}^1(r-\lambda_0)W_d(r)dr\end{bmatrix}+o_p(1).
\end{aligned}
$$

Proof of Theorem 6: Consider first the case with $d^* \in (-0.5, 0.5)$. After some algebra, we have

$$
\begin{aligned}
&M_T^*(T_1)\xrightarrow{p}M^*(\lambda)\\
&=\kappa^2(d)[\frac{\lambda+3}{\lambda}[W_d(1)]^2-\frac{6(\lambda+1)}{\lambda^2}W_d(1)\left(\int_0^1 rdW_d(r)\right)\\
&\quad+\frac{6}{\lambda^2(1-\lambda)}W_d(1)\left(\int_\lambda^1(r-\lambda)dW_d(r)\right)-\frac{6(2\lambda+1)}{\lambda^3(1-\lambda)}\left(\int_0^1 rdW_d(r)\right)\left(\int_\lambda^1(r-\lambda)dW_d(r)\right)\\
&\quad+\frac{3(3\lambda+1)}{\lambda^3}\left(\int_0^1 rdW_d(r)\right)^2+\frac{3}{\lambda^3(1-\lambda)^3}\left(\int_\lambda^1(r-\lambda)dW_d(r)\right)^2].
\end{aligned}
$$

We can write $\int_\lambda^1 (r-\lambda)dW_d(r) = -\lambda W_d(1) + \lambda W_d(\lambda) + \int_\lambda^1 r dW_d(r)$. Then

$$
\begin{aligned}
&M^*(\lambda) \\
&= \kappa^2(d)[\frac{\lambda+3}{\lambda}[W_d(1)]^2 - \frac{6(\lambda+1)}{\lambda^2} W_d(1) \left(\int_0^1 r dW_d(r)\right) \\
&- \frac{6}{\lambda(1-\lambda)^{1/2}} W_d(1) \left(\frac{W_d(1)-W_d(\lambda)}{\sqrt{1-\lambda}}\right) + \frac{6}{\lambda^2(1-\lambda)} W_d(1) \left(\int_\lambda^1 r dW_d(r)\right) \\
&+ \frac{6(2\lambda+1)}{\lambda^2(1-\lambda)^{1/2}} \left(\int_0^1 r dW_d(r)\right) \left(\frac{W_d(1)-W_d(\lambda)}{\sqrt{1-\lambda}}\right) \\
&- \frac{6(2\lambda+1)}{\lambda^3(1-\lambda)} \left(\int_0^1 r dW_d(r)\right) \left(\int_\lambda^1 r dW_d(r)\right) + \frac{3(3\lambda+1)}{\lambda^3} \left(\int_0^1 r dW_d(r)\right)^2 \\
&+ \frac{3}{\lambda^3(1-\lambda)^3} \left(-\lambda W_d(1) + \lambda W_d(\lambda) + \int_\lambda^1 r dW_d(r)\right)^2].
\end{aligned}
$$

Since the last term is quadratic, it dominates the other terms if it diverges. Note that

$$
\begin{aligned}
&\frac{3}{\lambda^3(1-\lambda)^3} \left(-\lambda W_d(1) + \lambda W_d(\lambda) + \int_\lambda^1 r dW_d(r)\right)^2 \\
&= \frac{3}{\lambda^3(1-\lambda)^2} \left(-\lambda \frac{W_d(1)-W_d(\lambda)}{\sqrt{1-\lambda}} + \frac{1}{(1-\lambda)^{1/2}} \int_\lambda^1 r dW_d(r)\right)^2. \qquad \text{(A.4)}
\end{aligned}
$$

By applying the law of iterated logarithms for a fractional Brownian motion, we can show that $\limsup_{\lambda\to 1} M^*(\lambda) = \infty$ *a.s.* for $d^* \in (-0.5, 0]$. Furthermore, note that

$$
\begin{aligned}
&\frac{3}{\lambda^3(1-\lambda)^3} \left(-\lambda W_d(1) + \lambda W_d(\lambda) + \int_\lambda^1 r dW_d(r)\right)^2 \\
&= \frac{3}{\lambda^3(1-\lambda)} \left(-\lambda \frac{W_d(1)-W_d(\lambda)}{1-\lambda} + \frac{1}{(1-\lambda)} \int_\lambda^1 r dW_d(r)\right)^2,
\end{aligned}
$$

Using the iterated law of logarithms for $d^* \in (0, 0.5)$, we obtain

$$
\limsup_{\lambda\to 1} \frac{W_d(1)-W_d(\lambda)}{1-\lambda} = \limsup_{s\to 0} \frac{W_d(s)}{s} = \infty \quad a.s..
$$

Hence, this shows that $\limsup_{\lambda\to 1} M^*(\lambda) = \infty$ *a.s.* for $d^* \in (0, 0.5)$. On the other hand, when $\lambda \to 0$, we can apply the law of iterated logarithms to the quadratic term only for $d^* \in (-0.5, 0]$ because the order of λ is not sufficient for the law of iterated logarithms to hold with $d^* \in (0, 0.5)$.

Proof of Theorem 7: Consider the case with $d^* \in (0.5, 1.5)$. With the functional central limit theorem in Lemma 3, it is easy to show that

$$
\begin{aligned}
& M_T^*(T_1) \xrightarrow{p} M^*(\lambda) \\
& = \kappa^2(d)[\frac{\lambda+3}{\lambda}\left(\int_0^1 W_d(r)dr\right)^2 - \frac{6(\lambda+1)}{\lambda^2}\left(\int_0^1 W_d(r)dr\right)\left(\int_0^1 rW_d(r)dr\right) \\
& \quad + \frac{6}{\lambda^2(1-\lambda)}\left(\int_0^1 W_d(r)dr\right)\left(\int_\lambda^1 (r-\lambda)W_d(r)dr\right) \\
& \quad - \frac{6(2\lambda+1)}{\lambda^3(1-\lambda)}\left(\int_0^1 rW_d(r)dr\right)\left(\int_\lambda^1 (r-\lambda)W_d(r)dr\right) + \frac{3(3\lambda+1)}{\lambda^3}\left(\int_0^1 rW_d(r)dr\right)^2 \\
& \quad + \frac{3}{\lambda^3(1-\lambda)^3}\left(\int_\lambda^1 (r-\lambda)W_d(r)dr\right)^2].
\end{aligned}
$$

Then, for any $\lambda \in (0,1)$,

$$
\begin{aligned}
M^*(\lambda) - M^*(0) &= M^*(\lambda) - M^*(1) \\
&= \kappa^2(d)[\sqrt{\frac{3(1-\lambda)}{\lambda}}\left(\int_0^1 W_d(r)dr\right) - \sqrt{\frac{3(3\lambda+1)}{\lambda^3} - 12}\left(\int_0^1 rW_d(r)dr\right) \\
&\quad + \sqrt{\frac{3}{\lambda^3(1-\lambda)^3}}\left(\int_\lambda^1 (r-\lambda)W_d(r)dr\right)]^2 > 0.
\end{aligned}
$$

The inequality holds because $M^*(\cdot)$ is not a constant process. This completes the proof.

References

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica, 59, 817-858.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point, Econometrica, 61, 821-856.

Andrews, D.W.K. and Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative, Econometrica, 62, 1383-1414.

Bai, J., 1994. Least squares estimation of a shift in linear processes, Journal of Time Series Analysis, 15, 453-472.

Bai, J., 1997. Estimation of a change point in multiple regressions, Review of Economics and Statistics, 79, 551-563.

Bai, J., 1998. A note on spurious break, Econometric Theory, 14, 663-669.

Bai, J., Lumsdaine, R.L. and Stock, J.H., 1998. Testing for and dating breaks in multivariate time series, Review of Economic Studies, 65, 395-432.

Bai, J. and Perron, P., 1998. Estimating and testing linear models with multiple structural changes, Econometrica, 66, 47-78.

Bai, J. and Perron, P., 2003. Computation and analysis of multiple structural change models, Journal of Applied Econometrics, 18, 1-22.

Deng, A. and Perron, P., 2006. A comparison of alternative asymptotic frameworks to analyze a structural change in a linear time trend, Econometrics Journal, 9, 423-447.

Dickey, D.A. and Fuller, W.A., 1975. Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association, 74, 427-431.

Feder, P.I., 1975. On asymptotic distribution theory in segmented regression problem: identified case, Annals of Statistics, 3, 49-83.

Gil-Alana, L.A., 2008. Fractional integration and structural breaks at unknown periods of time, Journal of Time Series Analysis, 29, 163-185.

Granger, C.W.J. and Joyeux, R., 1980. An introduction to long memory time series models and fractional differencing, Journal of Time Series Analysis, 1, 15-29.

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R., 2009. Simple, robust and powerful tests of the breaking trend hypothesis, Econometric Theory, 25, 995-1029.

Hatanaka, M. and Yamada, K., 1999. 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, Engle, R.F. and White, H. (Eds.), Oxford University Press.

Hosking, J., 1981. Fractional differencing, Biometrika, 68, 165-176.

Hsu, Y-C. and Kuan, C-M., 2008. Change-point estimation of nonstationary I(d) processes, Economic Letters, 98, 115-121.

Iacone, F., Leybourne, S.J. and Taylor, A.M.R., 2013. Testing for a break in trend when the order of integration is unknown, Journal of Econometrics, 176, 30-45.

Kim, D. and Perron, P., 2009. Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses, Journal of Econometrics, 148, 1-13.

Kuan, C-M. and Hsu, C-C., 1998. Change-point estimation of fractionally integrated processes, Journal of Time Series Analysis, 19, 693-708.

Lavielle, M. and Moulines, E., 2000. Least-squares estimation of an unknown number of shifts in a time series, Journal of Time Series Analysis, 20, 33-59.

Lobato, I.N. and Velasco, C., 2007. Efficient wald tests for fractional unit roots, Econometrica, 75, 575-589.

Marinucci, D. and Robinson, P.M., 1999. Alternative forms of fractional Brownian motion, Journal of Statistical Planning and Inference, 80, 111-122.

Nunes, L.C., Kuan, C-M. and Newbold, P., 1995. Spurious break, Econometric Theory, 11, 736-749.

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis, Econometrica, 57, 1361-1401.

Perron, P., 1991, A test for changes in a polynomial trend function for a dynamic time series. Research Memorandum No. 363, Econometric Research Program, Princeton University.

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.

Perron, P. and Yabu, T., 2009. Testing for shifts in trend with an integrated or stationary noise component, Journal of Business & Economic Statistics, 27, 369-396.

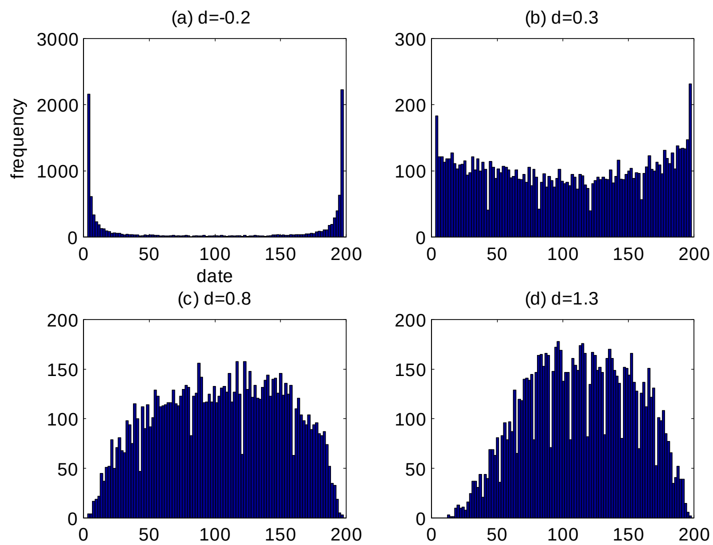
Perron, P. and Zhu, X., 2005. Structural breaks with deterministic and stochastic trends, Journal of Econometrics, 129, 65-119.

Taqqu, M., 1977. Law of the iterated logarithm for sums of non-linear functions of Gaussian random variables, Z. Wahr. Verw. Geb., 40, 203-238.

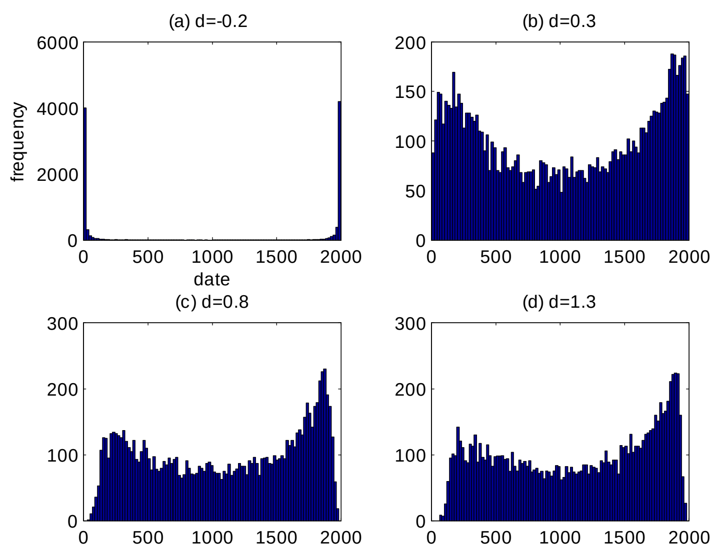
Vogelsang, T.J., 1997. Wald-type tests for detecting breaks in the trend function of a dynamic time series, Econometric Theory, 13, 818-849.

Vogelsang, T.J., 1999. Testing for a shift in trend when serial correlation is of unknown form, Manuscript, Department of Economics, Cornell University.

Wang, Q., Lin, Y-X. and Gulati, C.M., 2003. Asymptotics for general fractionally integrated processes with applications to unit root tests, Econometric Theory, 19, 143-164.

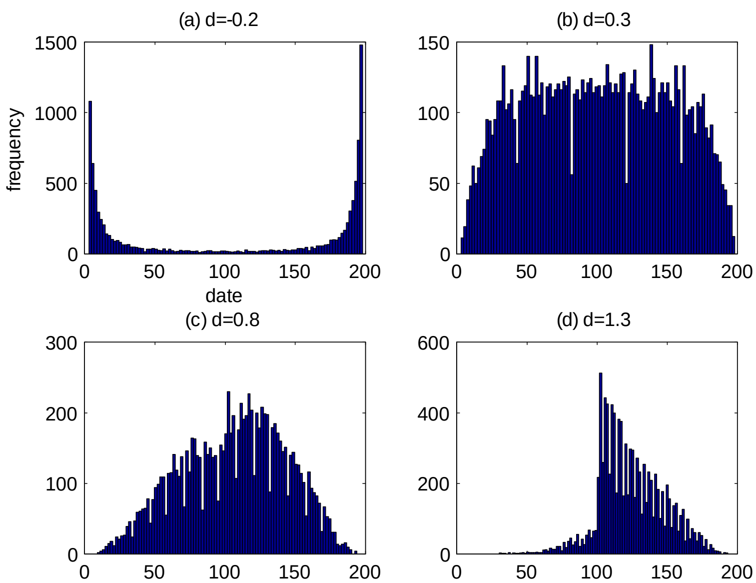


(a) T=200

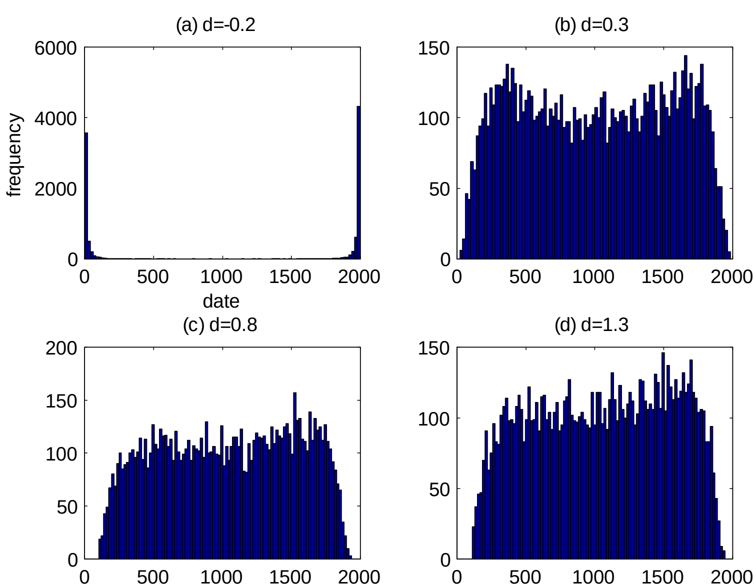


(b) T=2,000

Figure 1: Histograms of the estimate of a break date $\hat{T}_1$ for Model I



(a) T=200



(b) T=2,000

Figure 2: Histograms of the estimate of a break date $\hat{T}_1$ for Model II