Structural Change in AR(1) Models

Terence Tai-Leung Chong


Stable URL:
http://links.jstor.org/sici?sici=0266-4666%28200102%2917%3A1%3C87%3ASCIAM%3E2.0.CO%3B2-B

*Econometric Theory* is currently published by Cambridge University Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/cup.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Mon Mar 31 14:15:15 2008
This paper investigates the consistency of the least squares estimators and derives their limiting distributions in an AR(1) model with a single structural break of unknown timing. Let $\beta_1$ and $\beta_2$ be the preshift and postshift AR parameter, respectively. Three cases are considered: (i) $|\beta_1| < 1$ and $|\beta_2| < 1$; (ii) $|\beta_1| < 1$ and $\beta_2 = 1$; and (iii) $\beta_1 = 1$ and $|\beta_2| < 1$. Cases (ii) and (iii) are of particular interest but are rarely discussed in the literature. Surprising results are that, in both cases, regardless of the location of the change-point estimate, the unit root can always be consistently estimated and the residual sum of squares divided by the sample size converges to a discontinuous function of the change point. In case (iii), $\hat{\beta}_2$ does not converge to $\beta_2$ whenever the change-point estimate is lower than the true change point. Further, the limiting distribution of the break-point estimator for shrinking break is asymmetric for case (ii), whereas those for cases (i) and (iii) are symmetric. The appropriate shrinking rate is found to be different in all cases.

1. INTRODUCTION

Previous studies in the literature of structural change focus mainly on changes that take place in stationary processes, for example, a process that changes from one stationary process to another. Recently, there also have been studies on changes in nonstationary time series (Hansen, 1992). An interesting case of structural change is found in Mankiw and Miron (1986) and Mankiw, Miron, and Weil (1987); these authors conclude that the short term interest rate has changed from a stationary process to a near random walk since the Federal Reserve System was founded at the end of 1914. However, the asymptotic theory on this kind of structural change still remains unexplored.

In this paper, we develop a comprehensive asymptotic theory for an AR(1) model with a single structural break of unknown timing. Specifically, we examine the case where an AR(1) process changes from a stationary one to a nonstationary one (or the other way around). In each case, the asymptotic criterion function is derived, consistency of estimators is established, and limiting
distributions of estimators are also derived. As far as economic applications are concerned, we will restrict our discussion to the case where both the preshift and postshift parameters are in the interval \((-1,1]\).

The plan of the paper is as follows. Section 2 presents our basic model. Section 3 deals with the case where the preshift and postshift parameters are both less than one in absolute value. Section 4 studies the case where the preshift parameter is less than one in absolute value and the postshift parameter equals one. Section 5 examines the case where the preshift parameter equals one and the postshift parameter is less than one in absolute value. Section 6 discusses Monte Carlo experiments, and Section 7 concludes the paper. Mathematical details are collected in Appendices A–K.

Interesting findings include the following results: (i) in Sections 4 and 5 the asymptotic criterion function is random and has a sudden jump at the true break point; (ii) in Section 4, the asymptotic distribution of the break-point estimator for shrinking break is asymmetric; and (iii) in Section 5, the preshift estimator is always consistent, whereas the postshift estimator \(\hat{\beta}_2\) does not converge to \(\beta_2\) once the change-point estimate is lower than the true change point.

2. THE MODEL

Our basic model is an AR(1) model without drift, with a structural break in the AR parameter \(\beta\) at an unknown time \(k_0\). We consider the following model:

\[
y_t = \beta_1 y_{t-1} 1\{t \leq k_0\} + \beta_2 y_{t-1} 1\{t > k_0\} + \epsilon_t \quad (t = 1, 2, \ldots, T),
\]

where \(1\{\cdot\}\) is an indicator function that equals one when the statement in the braces is true and equals zero otherwise.

We let \(r_0 = k_0/T\) be the true break fraction and make the following assumptions.

(A1) \(y_0\) is drawn from an independent and identical distribution with zero mean and a finite variance.

(A2) \(\epsilon_t \sim i.i.d.(0, \sigma^2) \forall t, 0 < \sigma^2 < \infty, \) and \(E(\epsilon_t^4) < \infty.\)

(A3) \(\tau_0 \in [\tau, \tilde{\tau}] \subseteq (0,1).\)

We let \(k = \lfloor \tau T \rfloor\) where \(\lfloor \cdot \rfloor\) is the greatest integer function and define \(\hat{k} = \lfloor \tilde{\tau} T \rfloor\) and \(\tilde{k} = \lfloor \tilde{\tau} T \rfloor\).

Our interests are to estimate the structural parameters \(\beta_1\) and \(\beta_2\) and the time of change \(\tau_0\). We estimate the following model:

\[
\hat{y}_t = \hat{\beta}_1 y_{t-1} 1\{t \leq \lfloor \hat{\tau} T \rfloor\} + \hat{\beta}_2 y_{t-1} 1\{t > \lfloor \hat{\tau} T \rfloor\}.
\]
The estimation method employed here is the least squares method proposed in Bai (1994a, 1994b) and Chong (1995). For any given \( \tau \), the ordinary least squares (OLS) estimators are given by

\[
\hat{\beta}_1(\tau) = \frac{\sum_{t=1}^{[\tau T]} y_t y_{t-1}}{\sum_{t=1}^{[\tau T]} y_{t-1}^2},
\]

and the change-point estimator satisfies

\[
\hat{\tau} = \text{Arg min} \quad \text{RSS}_T(\tau),
\]

where

\[
\text{RSS}_T(\tau) = \sum_{t=1}^{[\tau T]} (y_t - \hat{\beta}_1(\tau)y_{t-1})^2 + \sum_{[\tau T]+1}^{T} (y_t - \hat{\beta}_2(\tau)y_{t-1})^2.
\]

We write \( \hat{k} = [\hat{\tau} T] \), \( \hat{\beta}_1 = \hat{\beta}_1(\hat{\tau}) \), and \( \hat{\beta}_2 = \hat{\beta}_2(\hat{\tau}) \). Throughout this paper, we let \( B_1(\cdot) \) and \( B_2(\cdot) \) be two independent Brownian motions defined on the non-negative half real line \( \mathbb{R}^+ \) and \( B(\cdot) \) and \( \tilde{B}(\cdot) \) be two independent standard Brownian motions on \([0, 1]\). The symbol \( \Rightarrow \) denotes the weak convergence of a stochastic process, \( \xrightarrow{P} \) represents convergence in probability, \( \xrightarrow{d} \) denotes convergence in distribution and \( \overset{d}{=} \) stands for identical in distribution. To achieve notational economy, the integral of a Brownian motion with respect to Lebesgue measure \( \int B(r) \, dr \) is written as \( \int B \). The stochastic integral \( \int B(r) \, dB(r) \) is written as \( \int B dB \).

### 3. CASE WHERE \( |\beta_1| < 1 \) AND \( |\beta_2| < 1 \)

#### 3.1. The Asymptotic Criterion Function

Consider the model in (1) with \( |\beta_1| < 1 \) and \( |\beta_2| < 1 \). A similar model has been studied by Salazar (1982), who provides a Bayesian analysis of structural changes in stationary AR(1) and AR(2) models with a known change point. In our model, the change point \( \tau_0 \) is unknown and has to be estimated by \( \hat{\tau} \). To show the consistency of \( \hat{\tau} \), the usual practice is to show that \( (1/T)\text{RSS}_T(\tau) \) converges uniformly to a nonstochastic function that has a unique minimum at \( \tau = \tau_0 \). Thus, we will focus on the asymptotic behavior of \( (1/T)\text{RSS}_T(\tau) \). The
following lemma is useful in deriving the limiting behavior of the criterion function \((1/T)RSSF(\tau)\) and in proving Theorem 1, which follows.

**LEMMA 1.** Let \(\{y_i\}_{i=1}^T\) be generated according to model (1) with \(|\beta_1| < 1\) and \(|\beta_2| < 1\). Under Assumptions (A1)-(A3), we have

\[
(3a) \quad \sup_{0 \leq \tau_0 \leq \tau \leq 1} \frac{1}{T} \sum_{i=1}^{\left\lfloor \tau T \right\rfloor} y_i - \beta \varepsilon_i = o_p(1);
\]

\[
(3b) \quad \frac{1}{T} \sum_{i=1}^{\left\lfloor \tau T \right\rfloor} y_i^2 \to \frac{\tau_0 \sigma^2}{1 - \beta_1^2};
\]

\[
(3c) \quad \frac{1}{T} \sum_{i=\left\lfloor \tau_0 T \right\rfloor+1}^T y_i^2 \to \frac{(1 - \tau_0) \sigma^2}{1 - \beta_2^2};
\]

\[
(3d) \quad \sup_{\tau = \tau_0} \left| \frac{1}{T} \sum_{i=\left\lfloor \tau T \right\rfloor+1}^{\left\lfloor \tau T \right\rfloor} y_i^2 - \frac{(\tau_0 - \tau) \sigma^2}{1 - \beta_1^2} \right| = o_p(1);
\]

\[
(3e) \quad \sup_{\tau = \tau_0} |\hat{\beta}_1(\tau) - \beta_1| = o_p(1);
\]

\[
(3f) \quad \sup_{\tau = \tau_0} \left| \frac{\hat{\beta}_2(\tau) - \frac{(\tau_0 - \tau)(1 - \beta_2^2)\beta_1 + (1 - \tau_0)(1 - \beta_1^2)\beta_2}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)}}{1 - \beta_2^2} \right| = o_p(1);
\]

\[
(3g) \quad \sup_{\tau_0 < \tau \leq \bar{T}} \left| \frac{1}{T} \sum_{i=\left\lfloor \tau T \right\rfloor+1}^{\left\lfloor \tau T \right\rfloor} y_i^2 - \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2} \right| = o_p(1);
\]

\[
(3h) \quad \sup_{\tau_0 < \tau < \bar{T}} \left| \frac{\hat{\beta}_1(\tau) - \frac{\tau_0(1 - \beta_2^2)\beta_1 + (\tau - \tau_0)(1 - \beta_1^2)\beta_2}{\tau_0(1 - \beta_2^2) + (\tau - \tau_0)(1 - \beta_1^2)}}{1 - \beta_1^2} \right| = o_p(1);
\]

\[
(3i) \quad \sup_{\tau_0 < \tau < \bar{T}} |\hat{\beta}_2(\tau) - \beta_2| = o_p(1).
\]

**Proof.** See Appendix A.

Note from (3e) that when \(\tau \leq \tau_0\), \(\hat{\beta}_1(\tau)\) converges uniformly to \(\beta_1\). This should be obvious because \(\hat{\beta}_1(\tau)\) only utilizes the data generated by the process \(y_i = \beta_1 y_{i-1} + \varepsilon_i\). However, from (3f), \(\hat{\beta}_2(\tau)\) converges uniformly to a weighted average of \(\beta_1\) and \(\beta_2\). The weight depends on the true change point, the true preshift and postshift parameters, and the location of \(\tau\). When \(\tau_0 < \tau \leq \bar{T}\), (3g) and (3i) display similar results.

From Appendix B, the asymptotic behavior of the residual sum of squares is as follows:

\[
\sup_{\tau = \tau_0} \left| \frac{1}{T} \text{RSS}_T(\tau) - \sigma^2 - \frac{(\tau_0 - \tau)(1 - \tau_0)(\beta_2 - \beta_1)^2 \sigma^2}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)} \right| = o_p(1),
\]

(6)

\[
\sup_{\tau_0 < \tau < \bar{T}} \left| \frac{1}{T} \text{RSS}_T(\tau) - \sigma^2 - \frac{\tau_0(\tau - \tau_0)(\beta_2 - \beta_1)^2 \sigma^2}{\tau_0(1 - \beta_2^2) + (\tau - \tau_0)(1 - \beta_1^2)} \right| = o_p(1).
\]

(7)
STRUCTURAL CHANGE IN AR(1) MODELS

One can easily verify that $(1/T) \text{RSS}_T(\tau)$ converges uniformly to a piecewise concave function of $\tau$ that attains a unique global minimum at $\tau = \tau_0$. A simulation of $(1/T) \text{RSS}_T(\tau)$ for $T = 4,000$ is plotted in Figure 1.

3.2. The Consistency and the Limiting Distributions of $\hat{\beta}_1$ and $\hat{\beta}_2$

If both $\beta_1$ and $\beta_2$ are less than one in absolute value then the process is said to be stationary throughout. It is not difficult to show that all the OLS estimators are consistent in this case. Bai (1994a, 1994b) shows that, in the conventional stationary case, the change-point estimator is $T$-consistent. This convergence rate is fast enough to make the limiting distributions of $\hat{\beta}_1$ and $\hat{\beta}_2$ behave as if the true change point $\tau_0$ is known. Theorem 1 establishes the asymptotic normality of $\hat{\beta}_1$ and $\hat{\beta}_2$.

**THEOREM 1.** Under assumptions (A1)–(A3), if $|\beta_1| < 1$ and $|\beta_2| < 1$, the OLS estimators $\hat{\tau}_T$, $\hat{\beta}_1(\hat{\tau}_T)$, and $\hat{\beta}_2(\hat{\tau}_T)$ are consistent and
Theorem. See Appendix C.

Thus, \( D_1(\tau_0) \) and \( D_2(\mu T) \) are asymptotically normally distributed, with variances depending on \( \beta_1, \beta_2, \) and \( \tau_0 \).

### 3.3. The Limiting Distribution of \( \hat{\tau}_T \) as \( \beta_2 \) Collapses to \( \beta_1 \)

In this section, we consider the limiting distribution of \( \hat{\tau}_T \). For a shift with fixed magnitude, Hinkley (1970) showed that the limiting distribution of the change-point estimator depends on the underlying distribution of the innovation \( \epsilon_i \) in a complicated manner. Thus, statistical inference on the change point under a break of fixed magnitude is difficult to perform. To obtain a limiting distribution of \( \hat{\tau}_T \) invariant to \( \epsilon_i \), we have to let the magnitude of change go to zero at an appropriate rate. Further, to ensure the consistency of \( \hat{\tau}_T \), we should let \( |\beta_2 - \beta_1| \) shrink at a rate slower than the rate of convergence of \( \hat{\beta}_1(\tau_0) \) and \( \hat{\beta}_2(\tau_0) \) so that there still exists a relative shift in parameters. To achieve this, we fix \( \beta_1 \) and let \( \beta_2T \) be a sequence of \( \beta_2 \) such that \( |\beta_2T| \to \frac{\beta_1}{\epsilon} \) and \( |\beta_2T| \to 1 - \beta_1 \) as \( T \to \infty \).

Recall from equation (5) that the change-point estimator is defined as

\[
\hat{\tau}_T = \arg\min_{\tau} \text{RSS}_T(\tau) = \arg\min_{\tau} \{\text{RSS}_T(\tau) - \text{RSS}_T(\tau_0)\}.
\]

Let \( v \in \mathbb{R} \) be a finite constant. We first examine the asymptotic behavior of \( \text{RSS}_T(\tau) - \text{RSS}_T(\tau_0) \) at the region \( \tau = \tau_0 + v/T(\beta_2T - \beta_1)^2 \). From Appendix D, we have the following expressions.

For \( v \leq 0 \),

\[
\text{RSS}_T(\tau) - \text{RSS}_T(\tau_0) \Rightarrow -2\sigma^2 B_1\left(\frac{|v|}{1 - \beta_1^2}\right) + \frac{\sigma^2 |v|}{1 - \beta_1^2}. \tag{9}
\]

For \( v > 0 \),

\[
\text{RSS}_T(\tau) - \text{RSS}_T(\tau_0) \Rightarrow -2\sigma^2 B_2\left(\frac{v}{1 - \beta_1^2}\right) + \frac{\sigma^2 v}{1 - \beta_1^2}. \tag{10}
\]

where \( B_1(\cdot) \) and \( B_2(\cdot) \) are defined in Section 2.
Let \( r = v/(1 - \beta_1^2) \) and applying the continuous mapping theorem for argmax functionals (Kim and Pollard, 1990), we have the following theorem.

**THEOREM 2.** Under assumptions (A1)–(A3), suppose we fix \( |\beta_1| < 1 \) and let \( \beta_{2T} \) be a sequence of \( \beta_2 \) such that \( |\beta_{2T} - \beta_1| \to 0 \) and \( \sqrt{T} |\beta_{2T} - \beta_1| \to \infty \) as \( T \to \infty \). Then the limiting distribution of \( \hat{\tau}_T \) is given by

\[
\frac{T(\beta_{2T} - \beta_1)^2}{1 - \beta_1^2} (\hat{\tau}_T - \tau_0) \xrightarrow{d} \text{Arg max } \left( B^*(r) - \frac{1}{2} |r| \right),
\]

where \( B^*(r) \) is a two-sided Brownian motion on \( R \) defined to be \( B_1(-r) \) for \( r < 0 \) and \( B_2(r) \) for \( r \geq 0 \). The terms \( B_1(\cdot) \) and \( B_2(\cdot) \) are defined in Section 2.

Proof. See Appendix D.

The exact distribution of the right hand side in (11) was first derived by Picard (1985). Yao (1987) tabulates the numerical approximation of this distribution. To understand the implication of Theorem 2, consider a Brownian motion moving along an inverted-V shape linear function with a kink at \( \tau_0 \). Because \( \hat{\tau}_T \) is the location where this motion achieves its maximum, its limiting distribution will be symmetric.

One should be careful that (11) is not the exact distribution of the changepoint estimator for fixed shifts as we are letting the shift shrink to zero when deriving the theorem. Rather, the theorem serves to provide a conservative confidence interval for \( \tau_0 \) when the shift is small.

### 4. CASE WHERE \( |\beta_1| < 1 \) and \( \beta_2 = 1 \)

We now consider the case where the AR(1) process shifts from a stationary process to an \( I(1) \) process. It is well known that the distribution theory for the least squares estimator in an AR process with a unit root or near unit root is nonstandard. See, for example, White (1958); Dickey and Fuller (1979); Lai and Siegmund (1983); Ahtola and Tiao (1984); Chan and Wei (1987, 1988); Phillips (1987, 1988), and Perron (1996).

#### 4.1. The Asymptotic Criterion Function

In our case, when \( \beta_2 = 1 \), the criterion function \((1/T)\text{RSS}_T(\tau)\) behaves very differently. The following lemma is useful in deriving its limiting behavior and in proving Theorem 3, which follows.

**LEMMA 2.** Let \( \{y_t\}_{t=1}^T \) be generated according to model (1) with \( |\beta_1| < 1 \) and \( \beta_2 = 1 \). Define

\[
\Xi_+ = (\tau_0, \hat{\tau}] \cap \left\{ \tau = \tau_0 + \frac{j}{T} : j = 1, 2, \ldots, [(1 - \tau_0)T] \right\} \quad \text{and} \quad \lim_{T \to \infty} \frac{j}{\sqrt{T}} = \infty.
\]
Then under Assumptions (A1)–(A3), we have

\begin{align}
(4a) & \left| \frac{1}{T} \sum_{t=[\tau_0 T]}^{[\tau T]} y_{t-1} e_t \right| = o_p(1); \\
(4b) & \frac{1}{T} \sum_{t=[\tau_0 T]+1}^{T} y_{t-1} e_t \Rightarrow \frac{1}{2} \left( 1 - \tau_0 \right) \sigma^2 (B^2(1) - 1) = O_p(1); \\
(4c) & \frac{1}{T^2} \sum_{t=[\tau_0 T]+1}^{T} y_{t-1}^2 \Rightarrow (1 - \tau_0)^2 \sigma^2 \int_0^1 B^2 = O_p(1); \\
(4d) & \sup_{t=\tau_0}^{\tau \leq \tau_0} \frac{1}{T} \sum_{t=[\tau_0 T]}^{[\tau T]} y_{t-1} e_t = o_p(1); \\
(4e) & \sup_{0 \leq t \leq \tau} \frac{1}{T} \sum_{t=[\tau_0 T]}^{[\tau T]} y_{t-1}^2 - \frac{(\tau_0 - \tau)\sigma^2}{1 - \beta_1^2} = o_p(1); \\
(4f) & \sup_{\tau_0 \leq \tau} |\beta_1(\tau) - \beta_1| = o_p(1); \\
(4g) & \sup_{\tau_0 \leq \tau} |\beta_2(\tau) - 1| = O_p\left( \frac{1}{\sqrt{T}} \right); \\
(4h) & \sup_{\tau_0 \leq \tau} |\beta_2(\tau) - \beta_1| = O_p(1); \\
(4i) & \frac{1}{T} \sum_{t=[\tau_0 T]+1}^{T} y_{t-1} e_t = o_p(1) \quad \text{for} \; \tau \in \Xi^T_+; \\
(4j) & \frac{1}{T} \sum_{t=[\tau_0 T]+1}^{T} y_{t-1} e_t = o_p(1) \quad \text{for} \; \tau \in \Xi^T_+; \\
(4k) & |\beta_1(\tau) - 1| = O_p\left( \frac{1}{\sqrt{T}} \right) \quad \text{for} \; \tau \in \Xi^T_+; \\
(4l) & |\beta_2(\tau) - 1| = O_p\left( \frac{1}{\sqrt{T}} \right) \quad \text{for} \; \tau \in \Xi^T_+. \\
\end{align}

Proof. See Appendix E.

Note that the set \( \Xi^T_+ \) is different from the set \((\tau_0, \bar{\tau})\) as \( \Xi^T_+ \) excludes sequences of \( \tau \) that converge too fast to \( \tau_0 \). For example, \( \Xi^T_+ \) contains all given constants \( \tau \) such that \( \tau_0 < \tau < \bar{\tau} \). The sequence \( \tau_T = \tau_0 + (1/\log T) \) also belongs to \( \Xi^T_+ \). However, the sequence \( \tau_T = \tau_0 + (1/T) \) is not in \( \Xi^T_+ \). From Appendix F, we have

\begin{equation}
\sup_{\tau_0 \leq \tau \leq \tau_0} \left| \frac{1}{T} \text{RSS}_T(\tau) - \sigma^2 - \frac{(1 - \beta_1)(\tau_0 - \tau)\sigma^2}{1 + \beta_1} \right| = o_p(1) \tag{12}
\end{equation}

and

\begin{equation}
\frac{1}{T} \text{RSS}_T(\tau) \xrightarrow{p} \sigma^2 + \frac{(1 - \beta_1)\tau_0\sigma^2}{1 + \beta_1} \quad \text{for} \; \tau \in \Xi^T_+. \tag{13}
\end{equation}
Note from (12) and (13) that \((1/T)RSS_T(\tau)\) converges uniformly to a downward sloping linear function for \(\tau \leq \tau \leq \tau_0\) and it converges to a flat line located above \(\sigma^2\) for \(\tau \in \Xi_T^r\). Further, there is an asymptotic gap between \((1/T)RSS_T(\tau_0)\) and \((1/T)RSS_T(\tau)\) for \(\tau \in \Xi_T^r\). The transitional process begins at \(\tau_0\) and completes at the left boundary of \(\Xi_T^r\). Thus the set \(\Xi_T^r\) establishes the rate at which \(\text{plim}(1/T)RSS_T(\tau)\) jumps from \(\sigma^2\) to a flat line located above \(\sigma^2\) once \(\tau > \tau_0\). This rate serves as a conservative rate of consistency of \(\hat{\tau}_T\).

To examine the consistency of \(\hat{\beta}_1\), we have to investigate the transitional behavior of \((1/T)RSS_T(\tau)\). It should be noted that, for any constant \(c > 0\),

\[
\hat{\beta}_1(\tau_0 + cT^{a-1}) = \theta_T(a, c) \left( \frac{\sum_{1}^{k_0} \varepsilon_i y_{t-1}}{\sum_{1}^{k_0} y_{t-1}} \right) + (1 - \theta_T(a, c)) \left( 1 + \frac{\sum_{k_0+1}^{k_0+[cT^a]} \varepsilon_i y_{t-1}}{\sum_{k_0+1}^{k_0+[cT^a]} y_{t-1}} \right),
\]

where \(\theta_T(a, c) = (\sum_{1}^{k_0} y_{t-1}^2)/((\sum_{1}^{k_0} y_{t-1}^2 + \sum_{k_0+1}^{k_0+[cT^a]} y_{t-1}^2))\).

For \(a < \frac{1}{2}\), we have

\[
\theta_T(a, c) \xrightarrow{p} 1, \quad \hat{\beta}_1(\tau_0 + cT^{a-1}) \xrightarrow{p} \beta_1,
\]

and

\[
\frac{1}{T} RSS_T(\tau_0 + cT^{a-1}) \xrightarrow{p} \sigma^2.
\]

For \(\frac{1}{2} < a < 1\), we have

\[
\theta_T(a, c) \xrightarrow{p} 0, \quad \hat{\beta}_1(\tau_0 + cT^{a-1}) \xrightarrow{p} 1,
\]

and

\[
\frac{1}{T} RSS_T(\tau_0 + cT^{a-1}) \xrightarrow{p} \sigma^2 + \frac{(1 - \beta_1)\tau_0\sigma^2}{1 + \beta_1}.
\]

This implies that if the convergence rate of \(\hat{\tau}_T\) is faster than \(T^{1/2}\) then \(\hat{\beta}_1\) will be consistent; otherwise it will be inconsistent.
Note that the transition is defined only when $a = \frac{1}{2}$, which means the speed of the transition is $1/\sqrt{T}$. For $a = \frac{1}{2}$, Appendix F shows that
\[
\theta_T\left(\frac{1}{2};c\right) \Rightarrow \frac{\tau_0}{\tau_0 + (1 - \beta_1^2) \int_0^c B_3^2},
\]
\[
\hat{\beta}_1\left(\tau_0 + \frac{c}{\sqrt{T}}\right) \Rightarrow \frac{\tau_0 \beta_1 + (1 - \beta_1^2) \int_0^c B_3^2}{\tau_0 + (1 - \beta_1^2) \int_0^c B_3^2},
\]
\[
\frac{1}{T} \text{RSS}_T\left(\tau_0 + \frac{c}{\sqrt{T}}\right) \Rightarrow \sigma^2 + \frac{\tau_0 \sigma^2 (1 - \beta_1^2)}{\tau_0 \left(\int_0^c B_3^2\right)^{-1} + 1 - \beta_1^2} \overset{\text{def}}{=} \sigma^2 + \varphi(c), \quad (14)
\]

where $B_3(\cdot)$ is a Brownian motion defined on $R_+$. Note that when $c = 0$, $(1/T)\text{RSS}_T(\tau_0) \xrightarrow{p} \sigma^2$, and as $c \to \infty$, $(1/T)\text{RSS}_T(\tau_0 + (c/\sqrt{T})) \xrightarrow{p} \sigma^2 + [(1 - \beta_1) \tau_0 \sigma^2]/(1 + \beta_1)$. Because $\partial \varphi(c)/\partial c > 0$, the transition is monotonically increasing.

Figure 2 simulates the behavior of $(1/T)\text{RSS}_T(\tau)$ using $T = 4,000$. For $\tau \leq \tau_0$, it is linear and decreasing. For $\tau > \tau_0$, it is flat and has a sudden jump near $\tau_0$.

4.2. The Consistency and the Limiting Distributions of $\hat{\tau}_T$, $\hat{\beta}_1$, and $\hat{\beta}_2$

For a fixed magnitude of the break, Theorem 3, which follows, shows that $\hat{\beta}_1$ is asymptotically normally distributed with a variance depending on $\beta_1$ and $\tau_0$, whereas $\hat{\beta}_2$ has a scaled Dickey–Fuller distribution.

Similar to Section 3.3, we should let the magnitude of the break go to zero at a certain rate to obtain the asymptotic distribution of $\hat{\tau}_T$. However, if we make $\beta_1$ converge to $\beta_2 (= 1)$, the process $\{y_t\}_{t=1}^{k_0}$ will be a near unit-root process. Thus, we are dealing with a structural change from a near unit-root process to an exact unit-root process.

We fix $\beta_2$ at 1 and let $\beta_{1T}$ be a sequence of $\beta_1$ such that $\beta_{1T} \to 1$ and $T(1 - \beta_{1T}) \to \infty$ as $T \to \infty$.

Consider the region $\tau = \tau_0 + v/T(1 - \beta_{1T})$ where $v \in R$. From Appendix G, we have the following expressions.

For $v \leq 0$,
\[
\text{RSS}_T(\tau) - \text{RSS}_T(\tau_0) \Rightarrow -2\sigma^2 B_1(|v|)B_\alpha(\frac{1}{2}) + |v|\sigma^2 B_\alpha^2(\frac{1}{2}).
\]
FIGURE 2. Graph of \((1/T)\text{RSS}_T(\tau)\) for \(\beta_1 = 0.5, \beta_2 = 1, T = 4000\).

For \(u > 0\),

\[
\text{RSS}_T(\tau) - \text{RSS}_T(\tau_0) = 2\sigma^2 \int_0^u \left( B_2(s) + B_a \left( \frac{1}{2} \right) \right) dB_2(s) + \sigma^2 \int_0^u \left( B_2(s) + B_a \left( \frac{1}{2} \right) \right)^2 ds,
\]

where \(B_1(\cdot)\) and \(B_2(\cdot)\) are defined in Section 2. The expression \(B_a(\frac{1}{2})\) is generated by \(\int_0^\infty \exp(-s) dB_1(s)\).

Theorem 3, which follows, states that as \(\beta_1\) approaches one, the limiting distribution of \(\hat{\tau}_T\) is the argmax of a random process moving along an inverted-V shape linear function on the real line.

**THEOREM 3.** *Under Assumptions (A1)–(A3), if \(|\beta_1| < 1\) and \(\beta_2 = 1\), the OLS estimators \(\hat{\tau}_T, \hat{\beta}_1(\hat{\tau}_T),\) and \(\hat{\beta}_2(\hat{\tau}_T)\) are all consistent and*
Suppose we fix $\beta_2$ at one and let $\beta_{1T}$ be a sequence of $\beta_1$ such that $|1 - \beta_{1T}| \to 0$ and $T(1 - \beta_{1T}) \to \infty$ as $T \to \infty$. Then the limiting distribution of $\hat{\tau}_T$ is given by

$$\sqrt{T}(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \xrightarrow{d} N\left(0, \frac{1 - \beta_1^2}{\tau_0}\right),$$

and

$$T(\hat{\beta}_2(\hat{\tau}_T) - 1) \Rightarrow \frac{B^2(1) - 1}{2(1 - \tau_0) \int_0^1 B^2}.$$ 

Proof. See Appendix G.

An interesting and important feature of Theorem 3 is that the distribution of $\hat{\tau}_T$ is asymmetric about $\tau_0$. An intuitive explanation for this result is that $(1/T)RSS_T(\tau)$ for breaks with fixed magnitude is asymmetric in the neighborhood of $\tau_0$ as shown in Figure 2. As we compress the size of break, this asymmetry perseveres. Thus, $\hat{\tau}_T = \arg \min \text{RSS}_T(\tau)$ should also be asymmetric about $\tau_0$.

5. CASE WHERE $\beta_1 = 1$ AND $|\beta_2| < 1$

Consider the case where $\beta_1 = 1$ and $|\beta_2| < 1$. Intuitively, it seems that the results of this case are the mirror images of the case where $|\beta_1| < 1$ and $\beta_2 = 1$. However, we show that a time reversing argument cannot be applied to model (1). The findings in this section turn out to be very different from those in Section 4.
5.1. The Asymptotic Criterion Function

The following lemma is useful in deriving the limiting behavior of \((1/T)\text{RSS}_T(\tau)\) and in proving Theorem 4, which follows.

**LEMMA 3.** Let \(\{y_t\}_{t=1}^T\) be generated according to model (1) with \(\beta_1 = 1\) and \(|\beta_2| < 1\). Define

\[
\Xi^T = \left[\tau, \tau_0\right] \cap \left\{ \tau = \tau_0 - \frac{j}{T}: j = 1, 2, \ldots, \lfloor \tau_0 T \rfloor \quad \text{and} \quad \lim_{T \to \infty} \frac{j}{\sqrt{T}} = \infty \right\}
\]

and let \(\Xi^T_+\) be defined as in Lemma 2. Under assumptions (A1)–(A3), we have

\begin{align*}
(5a) \quad & \frac{1}{T} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \Rightarrow \frac{\sigma^2}{2} (B^2(\tau_0) - \tau_0) = O_p(1); \\
(5b) \quad & \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \Rightarrow N\left(0, \frac{\sigma^4}{1 - \beta_2^2}\right); \\
(5c) \quad & \frac{1}{T^2} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \Rightarrow \frac{\tau - \tau_0 + B^2(\tau_0)}{1 - \beta_2^2} \sigma^2; \\
(5d) \quad & \sup_{\tau \in \Xi^T_+} \left| \frac{1}{T} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \right| = O_p(1); \\
(5e) \quad & \sup_{\tau \in \Xi^T_+} \left| \frac{1}{T} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \right| = O_p(1); \\
(5f) \quad & \sup_{\tau \in \Xi^T_+} \left| \frac{1}{T} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \right| = O_p(1); \\
(5g) \quad & \sup_{\tau \in \Xi^T_+} \left| \hat{\beta}_1(\tau) - 1 \right| = O_p\left(\frac{1}{T}\right); \\
(5h) \quad & \sup_{\tau \in \Xi^T_+} \left| \hat{\beta}_2(\tau) - 1 \right| = o_p(1); \\
(5i) \quad & \sup_{\tau \in \Xi^T_+} \left| \frac{1}{T} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \right| = o_p(1); \\
(5j) \quad & \sup_{\tau \in \Xi^T_+} \left| \frac{1}{T} \sum_{t=1}^{\lfloor \tau_0 T \rfloor} y_{t-1} e_t \right| = o_p(1); \\
(5k) \quad & \sup_{\tau \in \Xi^T_+} \left| \hat{\beta}_1(\tau) - 1 \right| = O_p\left(\frac{1}{T}\right); \\
(5l) \quad & \sup_{\tau \in \Xi^T_+} \left| \hat{\beta}_2(\tau) - \beta_2 \right| = o_p(1).
\end{align*}
Proof. See Appendix H.

The set $\Xi^T_\tau$ serves a similar purpose to that of $\Xi^T_+$ in Lemma 2. The term $\Xi^T_\tau$ excludes sequences of $\tau$ that converge to $\tau_0$ at a speed faster than $1/\sqrt{T}$. For example, $\Xi^T_\tau$ contains all given constants $\tau$ such that $\tau \leq \tau < \tau_0$. The sequence $\tau_T = \tau_0 - (1/\log T)$ also belongs to $\Xi^T_\tau$, but the sequence $\tau_T = \tau_0 - (1/T)$ does not.

We now derive the asymptotic behavior of $(1/T)RSS_T(\tau)$. From Appendix I, we have the following expressions.

For $\tau \in \Xi^T_\tau$,

$$
\frac{1}{T} RSS_T(\tau) \Rightarrow \sigma^2 + \frac{(1 - \beta_2)(1 - \tau_0 + B^2(\tau_0))}{1 + \beta_2} \sigma^2,
$$

$$(16)$$

For $\tau = \tau_0$,

$$
\frac{1}{T} RSS_T(\tau_0) \overset{p}{\rightarrow} \sigma^2.
$$

(17)

For $\tau \in \Xi^T_\tau$,

$$
\frac{1}{T} RSS_T(\tau) \Rightarrow \sigma^2 + \frac{(1 - \beta_2)(\tau - \tau_0 + B^2(\tau_0))}{1 + \beta_2} \sigma^2,
$$

$$(18)$$

Note from (16) that, for $\tau \in \Xi^T_\tau$, $(1/T)RSS_T(\tau)$ converges to a flat line located randomly above $\sigma^2 + [(1 - \beta_2)(1 - \tau_0)\sigma^2]/(1 + \beta_2) > \sigma^2$.

For $\tau \in \Xi^T_\tau$, $(1/T)RSS_T(\tau)$ converges to a random upward-sloping linear function of $\tau$. The lower support for this line is $\sigma^2 + [(1 - \beta_2)(\tau - \tau_0)\sigma^2]/(1 + \beta_2) > \sigma^2$.

In both situations, $(1/T)RSS_T(\tau)$ converges to a random line whose position depends on the true change point $\tau_0$, the postshift AR parameter $\beta_2$, and the realization of the standard Brownian motion $B(\tau_0)$. Note that the larger the magnitude value of break $(1 - \beta_2)$, the more easily the change will be detected.

Note that we work with $\Xi^T_\tau$ and $\Xi^T_\tau$ in this section to reflect the fact that $\text{plim}(1/T)RSS_T(\tau)$ is discontinuous at both sides of $\tau_0$, whereas we work with the interval $[\tau, \tau_0]$ and $\Xi^T_\tau$ in Section 4 because $\text{plim}(1/T)RSS_T(\tau)$ is only discontinuous at the right-hand-side neighborhood of $\tau_0$. 
Figure 3 simulates the behavior of \((1/T)\text{RSS}_T(\tau)\) with \(T = 4,000\). For \(\tau \leq \tau_0\), it is a random flat line and has a sudden plunge near \(\tau_0\). For \(\tau > \tau_0\), it is linear, random, and increasing, which agrees with our theory.

5.2. The Monotonic Transition

Note that for any positive constant \(c\) bounded away from 0, there is an asymptotic gap between \((1/T)\text{RSS}_T(\tau_0)\) and \((1/T)\text{RSS}_T(\tau_0 \pm c)\). The speed of the transitional process is \(1/\sqrt{T}\), and the transition is completed in \(\Xi^L\) and \(\Xi^L_+\), respectively. To investigate the consistency of the change-point estimator, it is sufficient to show that the asymptotic transition of \((1/T)\text{RSS}_T(\tau)\) from \(\tau = \tau_0\) to \(\tau = \tau_0 \pm c\) is monotonically increasing.

**Figure 3.** Graph of \((1/T)\text{RSS}_T(\tau)\) for \(\beta_1 = 1, \beta_2 = 0.5, T = 4000\).
From Appendix J, for $m = 0, 1, 2, \ldots$ and $(m/T) \to 0$ as $T \to \infty$, we have
\[
\sup_m \left| \hat{\beta}_1 \left( \tau_0 - \frac{m}{T} \right) - 1 \right| = o_p(1)
\]
and
\[
\hat{\beta}_2 \left( \tau_0 - \frac{m}{T} \right) - \beta_2 \Rightarrow \frac{(1 - \beta_2^2)mb^2(\tau_0)(1 - \beta_2)}{(1 - \beta_2^2)mb^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)}, \quad (19)
\]
\[
\frac{1}{T} RSS_T \left( \tau_0 - \frac{m}{T} \right) \Rightarrow \sigma^2 + \frac{(1 - \tau_0 + B^2(\tau_0))(1 - \beta_2)^2mb^2(\tau_0)}{(1 - \beta_2^2)mb^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)} \sigma^2 \Rightarrow \frac{\sigma^2}{\sigma^2 + h_1(m)}. \quad (20)
\]

If $m = 0$, we have $(1/T)RSS_T(\tau_0) \xrightarrow{p} \sigma^2$. As
\[
m \to \infty, \quad \frac{1}{T} RSS_T \left( \tau_0 - \frac{m}{T} \right) \Rightarrow \sigma^2 + \frac{(1 - \beta_2)(1 - \tau_0 + B^2(\tau_0))\sigma^2}{1 + \beta_2}.
\]
Thus, we bridge the gap between $(1/T)RSS_T(\tau_0)$ and $(1/T)RSS_T(\tau_0 - c)$. For $\tau = \tau_0 + (m/T)$, $m = 0, 1, 2, \ldots$, and $(m/T) \to 0$ as $T \to \infty$, we have
\[
\sup_m \left| \hat{\beta}_1 \left( \tau_0 + \frac{m}{T} \right) - 1 \right| = o_p(1),
\]
\[
\sup_m \left| \hat{\beta}_2 \left( \tau_0 + \frac{m}{T} \right) - \beta_2 \right| = o_p(1).
\]
From Appendix J, we also have
\[
\frac{1}{T} RSS_T \left( \tau_0 + \frac{m}{T} \right) \Rightarrow \sigma^2 + \frac{(1 - \beta_2)B^2(\tau_0)(1 - \beta_2^2m)}{1 + \beta_2} \sigma^2 \Rightarrow \frac{\sigma^2}{\sigma^2 + h_2(m)}. \quad (21)
\]
If $m = 0$, we have $(1/T)RSS_T(\tau_0) \xrightarrow{p} \sigma^2$. As $m \to \infty$, $(1/T)RSS_T(\tau_0 + (m/T)) \Rightarrow \sigma^2 + \{(1 - \beta_2)B^2(\tau_0)/(1 + \beta_2)\} \sigma^2$. Thus, we bridge the gap between $(1/T)RSS_T(\tau_0)$ and $(1/T)RSS_T(\tau_0 + c)$.

Note from equations (20) and (21) that, for all $m > 0$, we have $h_1(m + 1) > h_1(m) > h_1(0) = 0$ and $h_2(m + 1) > h_2(m) > h_2(0) = 0$. Thus the transition is monotonic. Note also from equation (19) that, unless $m = 0$, $\hat{\beta}_2(\tau_0 - (m/T))$ does not converge to $\beta_2$!

Remarks. However, because of the special behavior of $(1/T)RSS_T(\tau)$ described in equations (16)–(18) and (20)–(21), we will have $\lim_{T \to \infty} \Pr(\hat{k} \neq k_0) = 0$. Thus, $\hat{\beta}_2$ should be a consistent estimator of $\beta_2$ in practice. See Theorem 4, which follows.
5.3. The Consistency and the Limiting Distributions of $\hat{\tau}_T$, $\hat{\beta}_1$, and $\hat{\beta}_2$

Despite the fact that $\hat{B}_2$ is extremely sensitive to the local behavior of $\hat{k}$, Appendix K shows that for a fixed magnitude of break, $\lim_{T \to \infty} \Pr(\hat{k} \neq k_0) = 0$. This means $T(\hat{\tau}_T - \tau_0) \overset{p}{\to} 0$. Hence the distribution of $\hat{\tau}_T$ degenerates very fast for any fixed magnitude of break. However, if we allow $\beta_2$ to converge to $\beta_1(=1)$, the process $\{y_t\}_{t=k_0+1}^T$ will be a near unit-root process with an initial value drawn from an $I(1)$ process. Thus, we are dealing with structural changes from an exact unit-root process to a near unit-root process with a nonstationary initial value.

We fix $\beta_1$ at one and let $\beta_2T$ be a sequence of $\beta_2$ such that $\sqrt{T}(1 - \beta_2T) \to 0$ and $T^{3/4}(1 - \beta_2T) \to \infty$ as $T \to \infty$.

Consider the region $T = \tau_0 + v/(1 - \beta_2T)^2T^2$, where $v \in R$. From Appendix K, we have the following expressions.

For $v \leq 0$,

$$RSS_T(\tau) - RSS_T(\tau_0) = -2B(\tau_0)\sigma^2B_1(|v|) + |v|\sigma^2B^2(\tau_0).$$

For $v > 0$,

$$RSS_T(\tau) - RSS_T(\tau_0) = -2B(\tau_0)\sigma^2B_2(v) + v\sigma^2B^2(\tau_0),$$

where $B_1(\cdot)$, $B_2(\cdot)$, and $B(\cdot)$ are defined in Section 2.

**THEOREM 4.** Under Assumptions (A1)--(A3), if $\beta_1 = 1$ and $|\beta_2| < 1$ then $\hat{\tau}_T$, $\hat{\beta}_1(\hat{\tau}_T)$, and $\hat{\beta}_2(\hat{\tau}_T)$ are consistent, and

$$\Pr(\hat{k} \neq k_0) \to 0,$$

$$T(\hat{\beta}_1(\hat{\tau}_T) - 1) \Rightarrow \frac{B^2(\tau_0) - \tau_0}{2 \int_{\tau_0}^0 B^2},$$

$$\sqrt{T}(\hat{\beta}_2(\hat{\tau}_T) - \beta_2) \Rightarrow \frac{\sqrt{1 - \beta_2^2B(1)}}{1 - \beta_0 + B^2(\tau_0)}.$$
Proof. See Appendix K.

Thus $\hat{\beta}_1$ has a scaled Dickey–Fuller distribution. The term $\hat{\beta}_2$ can be described as being asymptotically normally distributed with a random variance. This is because

$$\sqrt{T} (\hat{\beta}_2 - \beta_2) = \sum_{[\tau_0 T]+1}^{T} \frac{\varepsilon_t y_{t-1}}{\sqrt{T}} + o_p(1),$$

the numerator follows a central limit theorem, and the denominator converges to a random variable by (5b) and (5d) in Lemma 3.

The rapid rate of convergence of $\hat{k}$ in Theorem 4 is very surprising. The rate in this case is much faster than those in the previous two cases. In previous cases, we only have $|\hat{k} - k_0| = O_p(1)$; that is, the estimation of $k_0$ is always subject to an error of order $O_p(1)$. Now, with $\beta_1 = 1$ and $|\beta_2| < 1$, $k_0$ can be precisely estimated asymptotically, despite the fact that $k_0 \to \infty$ as $T \to \infty$.

Further, as $\beta_2$ approaches one, the limiting distribution of $\hat{\tau}_T$ is the argmax of a random process moving along an inverted-V shape linear function.

6. MONTE CARLO EXPERIMENTS

For empirical applications, we perform the following experiments to see how well our asymptotic results match the small-sample properties of the estimators. In all experiments, the sample size is set at $T = 200$ and the number of replications is set at $N = 20,000$; $\{y_t\}_{t=1}^T$ is generated from model (1); $\{\varepsilon_t\}_{t=1}^T \sim \text{nid}(0,1)$ and $y_0 \sim \text{nid}(0,1)$ independent of $\{\varepsilon_t\}_{t=1}^T$. The true change point is set at $T_0 = 0.5$.

Experiment 1a. This experiment verifies equation (19), which predicts that when $D_1 = 1$ and $|\beta_2| < 1$, the postshift structural estimator $\hat{\beta}_2(\tau)$ does not converge to $\beta_2$ whenever $\tau < \tau_0$. We let $\beta_1 = 1$; $\beta_2 = 0.5, 0, -0.5$.

Let $\tilde{\beta}_2(\tau)$ be the mean of $\hat{\beta}_2(\tau)$ in 20,000 replications. We consider those $\tau$ in the neighborhood of $\tau_0(= 0.5)$.

Note that when $\beta_1 = 1$ and $|\beta_2| < 1$, the estimate of $\beta_2$ is inaccurate for $\tau < \tau_0$. The strange behavior of $\hat{\beta}_2(\tau)$ is highlighted by italic figures in Table 1. Note that $\hat{\beta}_2(\tau)$ is not close to $\beta_2$ even for $\tau = \tau_0 - (1/T) = 0.495$.

Experiment 1b. With $|\beta_1| < 1$, $\beta_2 = 1$, and $\tau$ close to $\tau_0$, we will show that both $\hat{\beta}_1(\tau)$ and $\hat{\beta}_2(\tau)$ are very close to the true parameters $\beta_1$ and $\beta_2$, respectively. Let $\beta_1 = 0.5, 0, -0.5; \beta_2 = 1$.

The results in Table 2 do not image those in Table 1. Here $\hat{\beta}_1(\tau)$ is very close to $\beta_1$. Note that in experiments 1a and 1b, the estimates of the unit root are close to one in all cases.
Table 1. Results for experiment 1a

<table>
<thead>
<tr>
<th>$\tau \setminus \hat{\bar{\beta}}$</th>
<th>$\hat{\beta}_1(\tau)$</th>
<th>$\hat{\beta}_2(\tau)$</th>
<th>$\hat{\beta}_1(\tau)$</th>
<th>$\hat{\beta}_2(\tau)$</th>
<th>$\hat{\beta}_1(\tau)$</th>
<th>$\hat{\beta}_2(\tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1 = 1$</td>
<td>$\beta_2 = 0.5$</td>
<td>$\beta_1 = 1$</td>
<td>$\beta_2 = 0$</td>
<td>$\beta_1 = 1$</td>
<td>$\beta_2 = -0.5$</td>
<td></td>
</tr>
<tr>
<td>0.480</td>
<td>0.9819</td>
<td>0.7041</td>
<td>0.9817</td>
<td>0.4703</td>
<td>0.9819</td>
<td>0.1350</td>
</tr>
<tr>
<td>0.485</td>
<td>0.9820</td>
<td>0.6773</td>
<td>0.9819</td>
<td>0.4163</td>
<td>0.9821</td>
<td>0.0541</td>
</tr>
<tr>
<td>0.490</td>
<td>0.9822</td>
<td>0.6401</td>
<td>0.9821</td>
<td>0.3413</td>
<td>0.9823</td>
<td>-0.0550</td>
</tr>
<tr>
<td>0.495</td>
<td>0.9824</td>
<td>0.5869</td>
<td>0.9822</td>
<td>0.2249</td>
<td>0.9825</td>
<td>-0.2160</td>
</tr>
<tr>
<td>0.500</td>
<td>0.9826</td>
<td>0.4930</td>
<td>0.9824</td>
<td>0.0009</td>
<td>0.9826</td>
<td>-0.4939</td>
</tr>
<tr>
<td>0.505</td>
<td>0.9728</td>
<td>0.4915</td>
<td>0.9628</td>
<td>0.0012</td>
<td>0.9529</td>
<td>-0.4923</td>
</tr>
<tr>
<td>0.510</td>
<td>0.9702</td>
<td>0.4903</td>
<td>0.9622</td>
<td>0.0013</td>
<td>0.9450</td>
<td>-0.4913</td>
</tr>
<tr>
<td>0.515</td>
<td>0.9693</td>
<td>0.4899</td>
<td>0.9617</td>
<td>0.0013</td>
<td>0.9423</td>
<td>-0.4909</td>
</tr>
<tr>
<td>0.520</td>
<td>0.9689</td>
<td>0.4898</td>
<td>0.9612</td>
<td>0.0013</td>
<td>0.9409</td>
<td>-0.4908</td>
</tr>
</tbody>
</table>

Experiment 2. This experiment simulates the distributions of $\hat{\beta}_1$ and $\hat{\beta}_2$.

Case 1: $\beta_1 = 0.5, \beta_2 = -0.5$;
Case 2: $\beta_1 = 0.5, \beta_2 = 1$;
Case 3: $\beta_1 = 1, \beta_2 = 0.5$.

Figures 4 to 6 display the results for Cases 1 to 3, respectively.
For Case 1, Theorem 1 states that both $\hat{\beta}_1$ and $\hat{\beta}_2$ are asymptotically normally distributed. The theorem is well supported by Figure 4 even for a sample size of 200.
For Case 2, Figure 5 shows that the small-sample distribution of $\hat{\beta}_1$ is approximately normal, whereas $\hat{\beta}_2$ appears to have a Dickey–Fuller distribution. Both results agree with the theorem.

Table 2. Results for experiment 1b

<table>
<thead>
<tr>
<th>$\tau \setminus \hat{\bar{\beta}}$</th>
<th>$\hat{\beta}_1(\tau)$</th>
<th>$\hat{\beta}_2(\tau)$</th>
<th>$\hat{\beta}_1(\tau)$</th>
<th>$\hat{\beta}_2(\tau)$</th>
<th>$\hat{\beta}_1(\tau)$</th>
<th>$\hat{\beta}_2(\tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1 = 0.5$</td>
<td>$\beta_2 = 1$</td>
<td>$\beta_1 = 0$</td>
<td>$\beta_2 = 1$</td>
<td>$\beta_1 = -0.5$</td>
<td>$\beta_2 = 1$</td>
<td></td>
</tr>
<tr>
<td>0.480</td>
<td>0.4899</td>
<td>0.9817</td>
<td>-0.0001</td>
<td>0.9810</td>
<td>-0.4902</td>
<td>0.9788</td>
</tr>
<tr>
<td>0.485</td>
<td>0.4900</td>
<td>0.9820</td>
<td>-0.0000</td>
<td>0.9815</td>
<td>-0.4903</td>
<td>0.9799</td>
</tr>
<tr>
<td>0.490</td>
<td>0.4900</td>
<td>0.9823</td>
<td>-0.0000</td>
<td>0.9820</td>
<td>-0.4904</td>
<td>0.9809</td>
</tr>
<tr>
<td>0.495</td>
<td>0.4901</td>
<td>0.9827</td>
<td>-0.0000</td>
<td>0.9825</td>
<td>-0.4905</td>
<td>0.9819</td>
</tr>
<tr>
<td>0.500</td>
<td>0.4902</td>
<td>0.9830</td>
<td>-0.0000</td>
<td>0.9830</td>
<td>-0.4905</td>
<td>0.9830</td>
</tr>
<tr>
<td>0.505</td>
<td>0.4956</td>
<td>0.9830</td>
<td>0.0109</td>
<td>0.9830</td>
<td>-0.4756</td>
<td>0.9830</td>
</tr>
<tr>
<td>0.510</td>
<td>0.5044</td>
<td>0.9830</td>
<td>0.0301</td>
<td>0.9830</td>
<td>-0.4512</td>
<td>0.9829</td>
</tr>
<tr>
<td>0.515</td>
<td>0.5158</td>
<td>0.9830</td>
<td>0.0559</td>
<td>0.9830</td>
<td>-0.4193</td>
<td>0.9829</td>
</tr>
<tr>
<td>0.520</td>
<td>0.5290</td>
<td>0.9829</td>
<td>0.0862</td>
<td>0.9830</td>
<td>-0.3819</td>
<td>0.9829</td>
</tr>
</tbody>
</table>
Figure 4. (a) Distribution of $\sqrt{\tau_0 T/(1 - \beta_1^2)}(\hat{\beta}_1 - \beta_1)$; --- ($T = 200$), ------ ($T = \infty$). (b) Distribution of $\sqrt{(1 - \tau_0) T/(1 - \beta_2^2)}(\hat{\beta}_2 - \beta_2)$; --- ($T = 200$), ------ ($T = \infty$).
Figure 5. (a) Distribution of $\sqrt{\tau_0 T/(1 - \beta_1^2)}(\hat{\beta}_1 - \beta_1)$; $\ldots \ldots (T = 200)$, $\ldots \ldots (T = \infty)$. (b) Distribution of $(1 - \tau_0)T/(\hat{\beta}_2 - \beta_2)$; $\ldots \ldots (T = 200)$, $\ldots \ldots (T = \infty)$. 
FIGURE 6. (a) Distribution of $\tau_0 T(\hat{\beta}_1 - \beta_1)$; $\cdots \cdots \cdot (T = 200)$, $\cdots \cdots \cdot (T = \infty)$. (b) Distribution of $\sqrt{T/(1 - \beta_2^2)} (\hat{\beta}_2 - \beta_2)$; $\cdots \cdots \cdot (T = 200)$, $\cdots \cdots \cdot (T = \infty)$. 
For Case 3, Theorem 4 predicts that $\hat{\beta}_1$ should have a Dickey–Fuller distribution and $\hat{\beta}_2$ should have a normal distribution. Figure 6 agrees with Theorem 4. Note that in a small sample, the variation of $T_T$ is nontrivial. This variation is the main source of the small-sample bias of $\hat{\beta}_1$ and $\hat{\beta}_2$. In Case 1 where both the pre- and postshift processes are stationary, the small-sample bias is very little. However, in Cases 2 and 3, $\hat{\beta}_1$ and $\hat{\beta}_2$ consist of both stationary and non-stationary observations; thus their small-sample distributions should look like a mixture of normal and Dickey–Fuller distributions. From Figures 5a and 6b, it is clear that the small-sample distribution has a larger variation and is more skew than a normal distribution. From Figures 5b and 6a, the small-sample distribution has a smaller variation than and is not as skew as a Dickey–Fuller distribution.

Experiment 3. In this experiment, we simulate the distribution of $T_T$ for shrinking break.

Case 1: $\beta_1 = 0.5$, $\beta_2T = 0.5 - T^{-1/3}$;
Case 2: $\beta_1T = 1 - T^{-1/4}$, $\beta_2 = 1$;
Case 3: $\beta_1 = 1$, $\beta_2T = 1 - T^{-5/8}$.

Thus, in each case, we simulate the behavior of $T^{3/4}(\hat{T} - \tau_0)$.

Figures 7(a)–7(c) display the results of Cases 1 to 3, respectively.

For Cases 1 and 3, Figures 7a and 7c show that $T_T$ is symmetrically distributed, as predicted by Theorems 1 and 4. For Case 2, Theorem 3 predicts that $T_T$ should be asymmetrically distributed. Figure 7b agrees with our predictions.

Theoretically, as $T \to \infty$, the domain for $T^{3/4}(\hat{T} - \tau_0)$ should be unbounded. However, because we fix $T$ at 200 in this experiment, and $T_T$ is searched within the (0,1) interval, hence the simulated distribution of $T^{3/4}(\hat{T} - \tau_0)$ is bounded in the interval $(-200^{3/4} \times 0.5, 200^{3/4} \times 0.5) = (-26.6, 26.6)$. Note that the distribution of $\hat{T}_T$ has a “boundary effect,” especially for Case 3, in that there is a lot of mass near the boundary. The clustering of mass at the boundary is due to the large value of $\text{Var}(\hat{\beta}_2(\tau) - \hat{\beta}_1(\tau))$ when $T$ nears zero or one. Although $|E(\hat{\beta}_2(\tau) - \hat{\beta}_1(\tau))| \leq |E(\hat{\beta}_2(\tau_0) - \hat{\beta}_1(\tau_0))|$ for all $\tau$, actual realization of $|\hat{\beta}_2(\tau) - \hat{\beta}_1(\tau)|$ may be larger than $|\hat{\beta}_2(\tau_0) - \hat{\beta}_1(\tau_0)|$ in finite sample, especially for the case where $\tau$ is near the boundary. For example, when $\tau$ is near 0, $\hat{\beta}_1(\tau)$ will be calculated based only on very few observations, and therefore it has a large variance. Note that $\hat{T}_T$ is generated by the minimization of $RSS_T(\tau)$, or in other words, by the maximization of $|\hat{\beta}_2(\tau) - \hat{\beta}_1(\tau)|$. Thus, the probability that $\hat{T}_T$ falls into the boundary is nontrivial. This phenomenon will disappear as the sample size goes to infinity.

7. CONCLUSION

AR models have been used extensively in economics. Many economic variables such as interest rates, real consumption, and real gross domestic product
(GDP) can be well predicted by their own lags. Some variables exhibit stationarity, whereas some variables display nonstationarity. A shock to the production technology or a sudden change in the government policy may cause structural changes in these autoregressive models. In particular, the changes may cause a stationary process to shift to a nonstationary one or vice versa.

In this paper, we present a structural change AR(1) model with independent and identically distributed (i.i.d.) innovations. We discuss three cases of struc-
tural changes. Case 1 deals with the change from a stationary process to another stationary process. Case 2 examines the situation where a stationary process shifts to an I(1) process. Case 3 discusses the situation where an I(1) process shifts to a stationary process with a nonstationary initial value. In each case, consistency of the least squares estimators is established, and their limiting distributions are derived. Having the limiting distributions of these estimates allows us to carry out statistical inference on the parameters of interest.

Cases 2 and 3 are intriguing. Results of these two cases do not mirror-image each other, which is counterintuitive. A possible explanation for this asymmetry is that for Case 2, the initial value of the postshift process is stationary, whereas Case 3 is nonstationary. Another explanation is that we are excluding an intercept in our model. As pointed out by Banerjee, Lumsdaine, and Stock (1992, p. 278), a more proper model for Case 3 should be, in our notations, $y_t = y_{t-1}1\{t \leq k_0\} + (y_{k_0}(1 - \beta_2) + \beta_2 y_{t-1})1\{t > k_0\} + \epsilon_t$. This process avoids a spurious sharp jump to zero at the break point.

It should also be mentioned that the disturbance terms $\{\epsilon_t\}_{t=1}^T$ in this paper are assumed to be i.i.d. In the case where both $|\beta_1|$ and $|\beta_2|$ are less than one, the assumption of i.i.d. avoids the inconsistency of $\hat{\beta}$’s due to serial correlation of $\epsilon_t$. This assumption is also very helpful in calculating the long run variances of processes such as $\sum \epsilon_t y_{t-1}$. However, extension to the cases of heterogeneous and/or dependent $\epsilon_t$ should be possible in the nonstationary cases at the expense of a more complicated mathematical treatment. Thus, a generalization of our model to ARMA($p,q$) models with an intercept deserves future investigation.

**REFERENCES**


**APPENDIX A:**

**PROOF OF LEMMA 1**

If $|\beta_1| < 1, |\beta_2| < 1, e_t$ are i.i.d., and $y_0$ is drawn from a stationary process, then

(1) (3a), (3d), and (3g) are consequences of the uniform law of large numbers in Andrews (1987, Theorem 1);

(2) (3b) and (3c) are consequences of the weak law of large numbers in Andrews (1988, Theorem 2).

\[
(3e) \quad \sup_{\tau \leq \tau \leq \tau_0} |\hat{\beta}_1(\tau) - \beta_1| = \sup_{\tau \leq \tau \leq \tau_0} \left| \frac{\sum_{t=1}^{[\tau]} y_{t-1} e_t}{\sum_{t=1}^{[\tau]} y_{t}^2} \right| \leq \frac{T}{\sum_{t=1}^{[\tau]} y_{t}^2} \sup_{\tau \leq \tau \leq \tau_0} \left| \frac{\sum_{t=1}^{[\tau]} y_{t-1} e_t}{T} \right| = O_P(1) o_p(1) = o_p(1).
\]
To show (3f), utilizing the fact that

$$\hat{\beta}_2(\tau) - \beta_1 = \frac{\sum_{[\tau_0]+1}^T y_{t-1}^2}{\sum_{[\tau]+1}^T y_{t-1}^2} (\beta_2 - \beta_1) + \frac{\sum_{[\tau]+1}^T e_t y_{t-1}}{\sum_{[\tau]+1}^T y_{t-1}^2},$$

adding and subtracting

$$T \frac{(1 - \tau_0)\sigma^2}{1 - \beta_2^2},$$

by the triangle inequality, and last by (3a), (3c), and (3d), we have

$$\sup_{T \leq \tau \leq \tau_0} \left| \hat{\beta}_2(\tau) - \frac{(\tau_0 - \tau)(1 - \beta_2^2)\beta_1 + (1 - \tau_0)(1 - \beta_1^2)\beta_2}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)} \right|$$

$$= \sup_{T \leq \tau \leq \tau_0} \left| \frac{\sum_{[\tau_0]+1}^T y_{t-1}^2}{\sum_{[\tau]+1}^T y_{t-1}^2} (\beta_2 - \beta_1) + \frac{\sum_{[\tau]+1}^T e_t y_{t-1}}{\sum_{[\tau]+1}^T y_{t-1}^2} \right|$$

$$- \frac{(1 - \tau_0)(1 - \beta_1^2)(\beta_2 - \beta_1)}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)}$$

$$\leq \sup_{T \leq \tau \leq \tau_0} \left| \frac{\sum_{[\tau]+1}^T e_t y_{t-1}}{\sum_{[\tau]+1}^T y_{t-1}^2} \right| + |\beta_2 - \beta_1| \sup_{T \leq \tau \leq \tau_0} \left| \frac{\sum_{[\tau]+1}^T y_{t-1}^2}{\sum_{[\tau]+1}^T y_{t-1}^2} - \frac{T (1 - \tau_0)\sigma^2}{1 - \beta_2^2} \right|$$

$$+ |\beta_2 - \beta_1| \sup_{T \leq \tau \leq \tau_0} \left| \frac{T (1 - \tau_0)\sigma^2}{1 - \beta_2^2} - \frac{(1 - \tau_0)(1 - \beta_2^2)}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)} \right|.$$
The proof for (3h) is analogous to (3f), using the fact that

\[ \hat{\beta}_1(\tau) - \beta_1 = \frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{i-1}^2}{\sum_{1}^{[\tau T]} y_{i-1}^2} \left( \beta_2 - \beta_1 \right) + \frac{\sum_{1}^{[\tau T]} e_i y_{i-1}}{\sum_{1}^{[\tau T]} y_{i-1}^2}; \]

adding and subtracting

\[ T \frac{(\tau - \tau_0)\sigma^2}{1 - \beta_2^2} \]

\[ (\beta_2 - \beta_1) \frac{\sum_{[\tau T]}^{[\tau_0 T]} y_{i-1}^2}{\sum_{1}^{[\tau T]} y_{i-1}^2}, \]
by the triangle inequality, and last by (3a), (3b), and (3g) we have

\[
\sup_{\tau_0 < \tau < \tau} \left| \hat{\beta}_1(\tau) - \frac{\tau_0 (1 - \beta_2^2) \beta_1 + (\tau - \tau_0) (1 - \beta_1^2) \beta_2}{\tau_0 (1 - \beta_2^2) + (\tau - \tau_0) (1 - \beta_1^2)} \right| 
\leq \sup_{\tau_0 < \tau < \tau} \left| \frac{\sum_{T} e_t y_{t-1}}{\sum_{T} y_{t-1}^2} \right| + \left| \beta_2 - \beta_1 \right| \sup_{\tau_0 < \tau < \tau} \left| \frac{T (\tau - \tau_0) \sigma^2}{1 - \beta_2^2} \right| \frac{\sum_{T} y_{t-1}^2}{\sum_{T} y_{t-1}^2}
\]

\[
+ \left| \beta_2 - \beta_1 \right| \sup_{\tau_0 < \tau < \tau} \left| \frac{T (\tau - \tau_0) \sigma^2}{1 - \beta_2^2} \right| \frac{\sum_{T} y_{t-1}^2}{\sum_{T} y_{t-1}^2} - \frac{(\tau - \tau_0) (1 - \beta_1^2)}{\tau_0 (1 - \beta_2^2) + (\tau - \tau_0) (1 - \beta_1^2)}
\]

\[
\leq \frac{T}{\sum_{T} y_{t-1}^2} \left| \frac{\sum_{T} e_t y_{t-1}}{T} \right| + \left| \beta_2 - \beta_1 \right| T \sup_{\tau_0 < \tau < \tau} \left| \frac{\sum_{T} y_{t-1}^2}{\sum_{T} y_{t-1}^2} \right| - \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2}
\]

\[
+ \left| \beta_2 - \beta_1 \right| (\tau - \tau_0) (1 - \beta_1^2) T \frac{\sum_{T} y_{t-1}^2}{\sum_{T} y_{t-1}^2}
\]

\[
\times \sup_{\tau_0 < \tau < \tau} \left| \frac{\tau_0 (1 - \beta_2^2) + (\tau - \tau_0) (1 - \beta_1^2)}{(1 - \beta_2^2)(1 - \beta_1^2)} \right| \sigma^2 \frac{\sum_{T} y_{t-1}^2}{T}
\]

\[
\leq \frac{T}{\sum_{T} y_{t-1}^2} \left| \frac{\sum_{T} e_t y_{t-1}}{T} \right| + \left| \beta_2 - \beta_1 \right| T \sup_{\tau_0 < \tau < \tau} \left| \frac{\sum_{T} y_{t-1}^2}{\sum_{T} y_{t-1}^2} \right| - \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2}
\]

\[
+ \left| \beta_2 - \beta_1 \right| T \frac{\sum_{T} y_{t-1}^2}{\sum_{T} y_{t-1}^2} + \sup_{\tau_0 < \tau < \tau} \left| \frac{\tau_0 \sigma^2}{1 - \beta_1^2} - \frac{\sum_{T} y_{t-1}^2}{T} \right| \left( \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2} - \frac{\sum_{T} y_{t-1}^2}{T} \right)
\]

\[= o_p(1). \]
These prove Lemma 1.

APPENDIX B: 
DERIVATION OF EQUATIONS (6) AND (7)

For $\tau \leq \tau \leq \tau_0$, we have

$$RSS_T(\tau) = \sum_{t=1}^{[\tau T]} (e_t - (\hat{\beta}_1(\tau) - \beta_1)y_{t-1})^2 + \sum_{t=[\tau T]+1}^{[\tau_0 T]} (e_t - (\hat{\beta}_2(\tau) - \beta_2)y_{t-1})^2$$

$$+ \sum_{t=[\tau_0 T]+1}^{T} (e_t - (\hat{\beta}_2(\tau) - \beta_2)y_{t-1})^2$$

$$= \sum_{t=1}^{T} e_t^2 - \left( \sum_{t=1}^{[\tau T]} y_{t-1} e_t \right)^2 - 2(\hat{\beta}_2(\tau) - \beta_1) \sum_{t=[\tau T]+1}^{[\tau_0 T]} y_{t-1} e_t$$

$$+ (\hat{\beta}_2(\tau) - \beta_1)^2 \sum_{t=[\tau_0 T]+1}^{[\tau_0 T]} y_{t-1}^2 - 2(\hat{\beta}_2(\tau) - \beta_2) \sum_{t=[\tau_0 T]+1}^{T} y_{t-1} e_t$$

$$+ (\hat{\beta}_2(\tau) - \beta_2)^2 \sum_{t=[\tau_0 T]+1}^{T} y_{t-1}^2.$$ 

(B.1)

By (B.1) and the fact that

$$\hat{\beta}_2(\tau) = \beta_1 \frac{\sum_{t=[\tau T]+1}^{[\tau T]} y_{t-1}^2}{\sum_{t=[\tau T]+1}^{[\tau T]} y_{t-1}^2} + \beta_2 \frac{\sum_{t=[\tau_0 T]+1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{t=[\tau_0 T]+1}^{[\tau_0 T]} y_{t-1}^2} + \frac{\sum_{t=[\tau T]+1}^{T} e_t y_{t-1}}{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2},$$
we have

$$RSS_{T}(\tau) - RSS_{T}(\tau_{0})$$

$$= \left( \frac{\sum_{1}^{[r_{0}T]} y_{t-1} \epsilon_{t}}{\sum_{1}^{[r_{0}T]} y_{t-1}^{2}} \right)^{2} - \left( \frac{\sum_{1}^{[rT]} y_{t-1} \epsilon_{t}}{\sum_{1}^{[rT]} y_{t-1}^{2}} \right)^{2}$$

$$- 2 \left( (\beta_{2} - \beta_{1}) \frac{\sum_{1}^{[r_{0}T]} y_{t-1}^{2}}{[r_{0}T]^{+1}} + \frac{\sum_{1}^{[rT]+1} \epsilon_{t} y_{t-1}}{[rT]+1} \right)\left( \frac{\sum_{1}^{[r_{0}T]} \epsilon_{t}}{[r_{0}T]^{+1}} \right)$$

$$+ \left( (\beta_{2} - \beta_{1}) \frac{\sum_{1}^{[r_{0}T]+1} y_{t-1}^{2}}{[r_{0}T]+1} + \frac{\sum_{1}^{[rT]+1} \epsilon_{t} y_{t-1}}{[rT]+1} \right)^{2}$$

$$- 2 \left( (\beta_{1} - \beta_{2}) \frac{\sum_{1}^{[r_{0}T]} y_{t-1}^{2}}{[r_{0}T]^{+1}} + \frac{\sum_{1}^{[rT]+1} \epsilon_{t} y_{t-1}}{[rT]+1} \right)\left( \frac{\sum_{1}^{[r_{0}T]} \epsilon_{t}}{[r_{0}T]^{+1}} \right)$$

$$+ \left[ (\beta_{1} - \beta_{2}) \frac{\sum_{1}^{[r_{0}T]} y_{t-1}^{2}}{[r_{0}T]^{+1}} + \frac{\sum_{1}^{[rT]+1} \epsilon_{t} y_{t-1}}{[rT]+1} \right]^{2} - \left( \frac{\sum_{1}^{[r_{0}T]} \epsilon_{t} y_{t-1}}{[r_{0}T]^{+1}} \right)^{2}$$

$$= 2(\beta_{2} - \beta_{1}) \left( \frac{\sum_{1}^{[r_{0}T]} y_{t-1}^{2}}{[r_{0}T]^{+1}} + \frac{\sum_{1}^{[rT]+1} \epsilon_{t} y_{t-1}}{[rT]+1} \right) - \frac{\sum_{1}^{[rT]} \epsilon_{t} y_{t-1}}{[rT]^{+1}}$$

$$+ (\beta_{2} - \beta_{1})^{2} \frac{\sum_{1}^{[r_{0}T]} y_{t-1}^{2}}{[r_{0}T]^{+1}} + \frac{\sum_{1}^{[rT]} y_{t-1}^{2}}{[rT]^{+1}} + \Lambda_{T}(\tau),$$

(B.2)

where

$$\Lambda_{T}(\tau) = \left( \frac{\sum_{1}^{[r_{0}T]} y_{t-1} \epsilon_{t}}{[r_{0}T]} \right)^{2} - \left( \frac{\sum_{1}^{[rT]} y_{t-1} \epsilon_{t}}{[rT]} \right)^{2} + \left( \frac{\sum_{1}^{[r_{0}T]+1} \epsilon_{t} y_{t-1}}{[r_{0}T]+1} \right)^{2} - \left( \frac{\sum_{1}^{[rT]+1} \epsilon_{t} y_{t-1}}{[rT]+1} \right)^{2}.$$
Let

\[ \lambda_1 = \left( \frac{(1 - \tau_0)(1 - \beta_1^2)(\beta_2 - \beta_1)}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)} \right)^2 \]

and

\[ \lambda_2 = \left( \frac{(\tau_0 - \tau)(1 - \beta_2^2)(\beta_2 - \beta_1)}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)} \right)^2. \]

By (B.1), the fact that

\[ \frac{(\beta_2 - \beta_1)^2 \sigma^2 (\tau_0 - \tau)(1 - \tau_0)}{(\tau_0 - \tau)(1 - \beta_2^2) + (1 - \tau_0)(1 - \beta_1^2)} = \lambda_1 \frac{(\tau_0 - \tau) \sigma^2}{1 - \beta_1^2} + \lambda_2 \frac{(1 - \tau_0) \sigma^2}{1 - \beta_2^2}, \]

and the triangle inequality, we have

\[
\sup_{\tau \geq \tau_0} \left| \frac{1}{T} \sum_{t=1}^{T} e_t^2 - \sigma^2 \right| + \sup_{\tau \geq \tau_0} \left| \frac{1}{T} \sum_{t=1}^{T} y_{t-1} \right|^2
\]

\[ \leq \sup_{\tau \geq \tau_0} \left| \frac{1}{T} \sum_{t=1}^{T} e_t^2 - \sigma^2 \right| + \sup_{\tau \geq \tau_0} \left| (\hat{\beta}_2(\tau) - \beta_1)^2 \frac{1}{T} \sum_{t=1}^{T} y_{t-1} e_t \right| \]

\[ + \sup_{\tau \geq \tau_0} \left| (\hat{\beta}_2(\tau) - \beta_1)^2 \frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \right| - \lambda_1 \frac{(\tau_0 - \tau) \sigma^2}{1 - \beta_1^2} \]

\[ + \sup_{\tau \geq \tau_0} \left| (\hat{\beta}_2(\tau) - \beta_2)^2 \frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \right| - \lambda_2 \frac{(1 - \tau_0) \sigma^2}{1 - \beta_2^2}. \]

We add and subtract \((\hat{\beta}_2(\tau) - \beta_1)^2[(\tau_0 - \tau) \sigma^2]/(1 - \beta_1^2)\) in the fourth term and \((\hat{\beta}_2(\tau) - \beta_2)^2[(1 - \tau_0) \sigma^2/(1 - \beta_2^2)]\) in the sixth term. By the triangle inequality, the fact that \(\sup|a^2 - b^2| \leq \sup|a - b|\sup|a + b|\), and Lemma 1, the preceding expression is bounded by
STRUCTURAL CHANGE IN AR(1) MODELS

\[
\begin{align*}
&\leq \frac{1}{T} \sum_{1}^{T} \epsilon_{t}^{2} - \sigma^{2} + \sup_{\tau \leq \tau \leq \tau_{0}} |\hat{\beta}_{1}(\tau) - \beta_{1}| \sup_{\tau \leq \tau \leq \tau_{0}} \frac{1}{T} \sum_{[\tau T]}^{[\tau T]} y_{i-1} e_{t} \\
&\quad + 2 \sup_{\tau \leq \tau \leq \tau_{0}} |\hat{\beta}_{2}(\tau) - \beta_{1}| \sup_{\tau \leq \tau \leq \tau_{0}} \frac{1}{T} \sum_{[\tau T] + 1}^{[\tau T]} y_{i-1} e_{t} \\
&\quad + \sup_{\tau \leq \tau \leq \tau_{0}} (\hat{\beta}_{2}(\tau) - \beta_{1})^{2} \sup_{\tau \leq \tau \leq \tau_{0}} \frac{1}{T} \sum_{[\tau T] + 1}^{[\tau T]} y_{i-1}^{2} - \frac{(\tau_{0} - \tau) \sigma^{2}}{1 - \beta_{1}^{2}} \\
&\quad + \sup_{\tau \leq \tau \leq \tau_{0}} \frac{(\tau_{0} - \tau) \sigma^{2}}{1 - \beta_{1}^{2}} \sup_{\tau \leq \tau \leq \tau_{0}} |(\hat{\beta}_{2}(\tau) - \beta_{1})^{2} - \lambda_{1}| \\
&\quad + 2 \sup_{\tau \leq \tau \leq \tau_{0}} |\hat{\beta}_{2}(\tau) - \beta_{2}| \left| \frac{\sum_{[\tau T] + 1}^{T} y_{i-1} e_{t}}{T} \right| \\
&\quad + \sup_{\tau \leq \tau \leq \tau_{0}} (\hat{\beta}_{2}(\tau) - \beta_{2})^{2} \left| \frac{\sum_{[\tau T] + 1}^{[\tau T]} y_{i-1}^{2}}{T} - \frac{(1 - \tau_{0}) \sigma^{2}}{1 - \beta_{2}^{2}} \right| \\
&\quad + \frac{(1 - \tau_{0}) \sigma^{2}}{1 - \beta_{2}^{2}} \sup_{\tau \leq \tau \leq \tau_{0}} |(\hat{\beta}_{2}(\tau) - \beta_{2})^{2} - \lambda_{2}| \\
&= o_{p}(1).
\end{align*}
\]

This derives equation (6).

For \( \tau_{0} < \tau \leq \tau \), we can write the residual sum of squares as follows:

\[
\text{RSS}_{\tau}(\tau) = \sum_{1}^{[\tau T]} (e_{t} - (\hat{\beta}_{1}(\tau) - \beta_{1}) y_{t-1})^{2} + \sum_{[\tau T] + 1}^{[\tau T]} (e_{t} - (\hat{\beta}_{1}(\tau) - \beta_{2}) y_{t-1})^{2} \\
+ \sum_{[\tau T] + 1}^{T} (e_{t} - (\hat{\beta}_{2}(\tau) - \beta_{2}) y_{t-1})^{2} \\
= \sum_{1}^{T} \epsilon_{t}^{2} - 2(\hat{\beta}_{1}(\tau) - \beta_{1}) \sum_{1}^{[\tau T]} y_{t-1} e_{t} + (\hat{\beta}_{1}(\tau) - \beta_{1})^{2} \sum_{1}^{[\tau T]} y_{t-1}^{2} \\
- 2(\hat{\beta}_{1}(\tau) - \beta_{2}) \sum_{[\tau T] + 1}^{[\tau T]} y_{t-1} e_{t} + (\hat{\beta}_{1}(\tau) - \beta_{2})^{2} \sum_{[\tau T] + 1}^{[\tau T]} y_{t-1}^{2} \\
= \left( \frac{\sum_{[\tau T] + 1}^{T} e_{t} y_{t-1}}{\sum_{[\tau T] + 1}^{T} y_{t-1}^{2}} \right)^{2}.
\]

(B.3)
By (B.3) and the fact that
\[
\hat{\beta}_1(\tau) = \sum_{t=1}^{[\tau \tau]} \frac{y_{t-1}}{\beta_1 + \sum_{t=1}^{[\tau \tau]} \frac{y_{t-1}^2}{\beta_2 + \sum_{t=1}^{[\tau \tau]} \frac{e_{t}y_{t-1}}{\sum_{t=1}^{[\tau \tau]} y_{t-1}^2}},
\]
we have
\[
RSS_T(\tau) - RSS_T(\tau_0) = 2(\beta_2 - \beta_1) \left( \sum_{t=1}^{[\tau \tau]} \frac{y_{t-1} e_{t}}{\sum_{t=1}^{[\tau \tau]} y_{t-1}^2} \right) - \frac{(\beta_2 - \beta_1)^2}{\sum_{t=1}^{[\tau \tau]} y_{t-1}^2} + \Lambda_T(\tau).
\]

Let
\[
\lambda_3 = \left( \frac{(\tau - \tau_0)(1 - \beta_1^2)(\beta_2 - \beta_1)}{\tau_0(1 - \beta_2^2) + (\tau - \tau_0)(1 - \beta_1^2)} \right)^2 \quad \text{and} \quad \lambda_4 = \left( \frac{\tau_0(1 - \beta_2^2)(\beta_2 - \beta_1)}{\tau_0(1 - \beta_2^2) + (\tau - \tau_0)(1 - \beta_1^2)} \right)^2.
\]

By (B.3), the fact that
\[
\frac{\tau_0(\beta_2 - \beta_1)^2 \sigma^2 (\tau - \tau_0)}{\tau_0(1 - \beta_2^2) + (\tau - \tau_0)(1 - \beta_1^2)} = \lambda_3 \frac{\tau_0 \sigma^2}{1 - \beta_1^2} + \lambda_4 \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2},
\]
and the triangle inequality, we have
\[
\sup_{\tau_0 < \tau \leq \tau} \left| \frac{1}{T} RSS_T(\tau) - \sigma^2 - \frac{\tau_0(\beta_2 - \beta_1)^2 \sigma^2 (\tau - \tau_0)}{\tau_0(1 - \beta_2^2) + (\tau - \tau_0)(1 - \beta_1^2)} \right| \\
= \frac{1}{T} \sum_{t=1}^{T} e_{t}^2 - \sigma^2 + \sup_{\tau_0 < \tau \leq \tau} \left| (\hat{\beta}_1(\tau) - \beta_1) \frac{1}{T} \sum_{t=1}^{[\tau \tau]} y_{t-1}^2 - \lambda_3 \frac{\tau_0 \sigma^2}{1 - \beta_1^2} \right| \\
+ \sup_{\tau_0 < \tau \leq \tau} \left| (\hat{\beta}_1(\tau) - \beta_1) \frac{2}{T} \sum_{t=1}^{[\tau \tau]} y_{t-1} e_{t} \right| \\
+ \sup_{\tau_0 < \tau \leq \tau} \left| (\hat{\beta}_1(\tau) - \beta_2) \frac{2}{T} \sum_{t=1}^{[\tau \tau]+1} y_{t-1} e_{t} \right| \\
+ \sup_{\tau_0 < \tau \leq \tau} \left| (\hat{\beta}_1(\tau) - \beta_2) \frac{2}{T} \sum_{t=1}^{[\tau \tau]+1} y_{t-1} \right| - \lambda_4 \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2} + \sup_{\tau_0 < \tau \leq \tau} \left( \frac{\sum_{t=1}^{[\tau \tau]+1} y_{t-1}^2}{T} \right)^2. 
\]
We add and subtract \((\hat{\beta}_1(\tau) - \beta_1)^2[\tau_0 \sigma^2/(1 - \beta_1^2)]\) in the third term and \((\hat{\beta}_1(\tau) - \beta_2)^2[(\tau - \tau_0) \sigma^2]/(1 - \beta_2^2)\) in the sixth term. Using \(\sup |a^2 - b^2| \leq \sup |a - b| \sup |a + b|, \sup (a - b)^2 \leq (\sup |a - b|)^2\), and Lemma 1, the preceding expression is bounded by

\[
\leq \left| \frac{1}{T} \sum_1^T e_t^2 - \sigma^2 \right| + \left| \frac{2}{T} \sum_1^{\tau_0 T} y_{t-1} e_t \right| \sup_{\tau_0 < \tau_t \leq \tau} |\hat{\beta}_1(\tau) - \beta_1| \\
+ \left| \frac{1}{T} \sum_1^{\tau_0 T} y_{t-1}^2 - \frac{\tau_0 \sigma^2}{1 - \beta_1^2} \right| \sup_{\tau_0 < \tau_t \leq \tau} (\hat{\beta}_1(\tau) - \beta_1)^2 \\
+ \frac{\tau_0 \sigma^2}{1 - \beta_1^2} \sup_{\tau_0 < \tau_t \leq \tau} |(\hat{\beta}_1(\tau) - \beta_1)^2 - \lambda_3| \\
+ 2 \sup_{\tau_0 < \tau_t \leq \tau} |\hat{\beta}_1(\tau) - \beta_2| \sup_{\tau_0 < \tau_t \leq \tau} \left| \frac{1}{T} \sum_1^{\tau_0 T+1} y_{t-1} e_t \right| \\
+ \sup_{\tau_0 < \tau_t \leq \tau} (\hat{\beta}_1(\tau) - \beta_2)^2 \sup_{\tau_0 < \tau_t \leq \tau} \left| \frac{1}{T} \sum_1^{\tau_0 T+1} y_{t-1}^2 - \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2} \right| \\
+ \sup_{\tau_0 < \tau_t \leq \tau} \frac{(\tau - \tau_0) \sigma^2}{1 - \beta_2^2} \sup_{\tau_0 < \tau_t \leq \tau} |(\hat{\beta}_1(\tau) - \beta_2)^2 - \lambda_4| \\
+ \sup_{\tau_0 < \tau_t \leq \tau} |\hat{\beta}_2(\tau) - \beta_2| \sup_{\tau_0 < \tau_t \leq \tau} \left| \frac{1}{T} \sum_1^T y_{t-1} e_t \right| \\
= o_p(1).
\]

This shows equation (7).

**APPENDIX C: PROOF OF THEOREM 1**

If \(|\hat{\tau}_T - \tau_0| = O_p(1/T)\), or equivalently \(|\hat{k} - k_0| = O_p(1)\), then for any \(\eta > 0\), there exists an \(M < \infty\) such that \(\Pr(|\hat{k} - k_0| > M) < \eta\). We shall prove this by using the contradiction argument. Suppose \(\hat{\tau}_T\) is not \(T\)-consistent; then there exists a sequence \(M_T > 0\) such that \(M_T \rightarrow \infty\), \((M_T/T) \rightarrow 0\) as \(T \rightarrow \infty\), and

\[
\lim_{T \rightarrow \infty} \Pr(|\hat{k} - k_0| > M_T) = \alpha,
\]

where \(\alpha\) is a positive constant in \((0,1]\).

Now, let \(\lambda = \beta_2 - \beta_1, Z_T = \{1,2,\ldots,T\}, D_{1T} = \{m: m \in Z_T, m < k_0 - M_T\}, D_{2T} = \{m: m \in Z_T, m > k_0 + M_T\}, D_{3T} = \{m: m \in Z_T, k_0 - M_T \leq m \leq k_0 + M_T\}.\)
Note that

\[
\Pr(|\hat{k} - k_0| > M_T) = \Pr\left(\inf_{m \in D_{1T} \cup D_{2T}} \text{RSS}_T\left(\frac{m}{T}\right) < \inf_{m \in D_{3T}} \text{RSS}_T\left(\frac{m}{T}\right)\right).
\]

Because \( k_0 \in D_{3T} \), we have \( \inf_{m \in D_{3T}} \text{RSS}_T(\frac{m}{T}) \leq \text{RSS}_T(\tau_0) \), and the preceding probability is bounded by

\[
\leq \Pr\left(\inf_{m \in D_{1T} \cup D_{2T}} \text{RSS}_T\left(\frac{m}{T}\right) < \text{RSS}_T(\tau_0)\right)
\]

\[
\leq \sum_{i=1}^{2} \Pr\left(\inf_{m \in D_{3T}} \left[ \text{RSS}_T\left(\frac{m}{T}\right) - \text{RSS}_T(\tau_0) \right] < 0\right).
\]

By using (B.2) and (B.4) in Appendix B, the right-hand side equals

\[
= \Pr\left(\inf_{m \in D_{1T}} \left\{ 2\lambda \left( \frac{\sum_{k_0+1}^{T} y_{T-1} e_t}{\sum_{k_0+1}^{T} y_{T-1}^2} - \frac{\sum_{m+1}^{k_0} y_{T-1} e_t}{\sum_{m+1}^{k_0} y_{T-1}^2} \right) + \lambda^2 + \frac{\sum_{m+1}^{k_0} y_{T-1}^2}{\sum_{k_0+1}^{T} \sum_{m+1}^{k_0} y_{T-1}^2} \Lambda_T\left(\frac{m}{T}\right) \right\} < 0 \right)
\]

\[
+ \Pr\left(\inf_{m \in D_{2T}} \left\{ 2\lambda \left( \frac{\sum_{k_0+1}^{m} y_{T-1} e_t}{\sum_{k_0+1}^{m} y_{T-1}^2} - \frac{\sum_{1}^{k_0} y_{T-1} e_t}{\sum_{1}^{k_0} y_{T-1}^2} \right) + \lambda^2 + \frac{\sum_{k_0+1}^{m} y_{T-1}^2}{\sum_{k_0+1}^{m} \sum_{1}^{k_0} y_{T-1}^2} \Lambda_T\left(\frac{m}{T}\right) \right\} < 0 \right).
\]

Using \(-\inf x = \sup(-x)\), \(\sup(x + y) \leq \sup x + \sup y\), \(\sup xy \leq \sup x \sup y\), and \(\sup x \leq \sup|x|\), the preceding item is bounded by

\[
\leq \Pr(-2\lambda A_1 + 2\lambda A_2 + A_3 > \lambda^2) + \Pr(2\lambda A_4 - 2\lambda A_5 + A_6 > \lambda^2),
\]
where

\[
A_1 = \frac{\sum_{k_0+1}^{T} y_{t-1} \varepsilon_t}{\sum_{k_0+1}^{T} y_{t-1}^2} = o_p(1)
\]

by (3a) and (3c) of Lemma 1;

\[
A_2 = \sup_{m \in D_1T} \frac{\sum_{k_0}^{m+1} y_{t-1} \varepsilon_t}{\sum_{k_0}^{m+1} y_{t-1}^2} \leq \sup_{m \in D_1T} \left| \frac{1}{k_0 - m} \sum_{m+1}^{k_0} y_{t-1} \varepsilon_t \right| = o_p(1)
\]

by the uniform law of large numbers in Andrews (1987, Theorem 1);

\[
A_3 = \sup_{m \in D_1T} \left| \frac{\sum_{m+1}^{T} y_{t-1}^2}{\sum_{k_0+1}^{m+1} y_{t-1}^2} \right| \leq \left( \frac{1}{T} \sum_{k_0+1}^{T} \varepsilon_t y_{t-1} \right) + \left( \frac{1}{T} \sum_{m+1}^{T} \varepsilon_t y_{t-1} \right)
\]

\[
\times \sup_{m \in D_1T} \left| \frac{\sum_{1}^{k_0} y_{t-1} \varepsilon_t}{\sum_{1}^{k_0} y_{t-1}^2} \right| \leq \left( \frac{1}{T} \sum_{k_0+1}^{T} \varepsilon_t y_{t-1} \right) + \left( \frac{1}{T} \sum_{m+1}^{T} \varepsilon_t y_{t-1} \right)
\]

\[
A_4 = \frac{\sum_{k_0+1}^{T} y_{t-1} \varepsilon_t}{\sum_{k_0+1}^{T} y_{t-1}^2} = o_p(1)
\]

by (3a) and (3b) of Lemma 1;

\[
A_5 = \sup_{m \in D_2T} \frac{\sum_{k_0+1}^{m} y_{t-1} \varepsilon_t}{\sum_{k_0+1}^{m} y_{t-1}^2} = o_p(1)
\]
by a similar argument as in $A_2$:

$$A_6 = \sup_{m \in D_T} \left| \frac{\sum_{l=1}^{m} y_{l-1}^2 - \sum_{l=1}^{k_0} y_{l-1}^2}{\sum_{l=1}^{k_0+1} y_{l-1}^2} A_T \left( \frac{m}{T} \right) \right| \leq \frac{1}{k_0} + \frac{1}{k_0 + M_T}$$

$$\times \left( \sup_{m \in D_T} \left| \frac{\left( \sum_{l=1}^{k_0} y_{l-1}^2 e_l \right)^2}{\sum_{l=1}^{k_0+1} y_{l-1}^2} - \left( \sum_{l=1}^{m} y_{l-1}^2 e_l \right)^2 \right| \right)$$

$$+ \sup_{m \in D_T} \left| \frac{\left( \sum_{l=k_0+1}^{T} e_l y_{l-1} \right)^2}{\sum_{l=k_0+1}^{T} y_{l-1}^2} - \left( \sum_{l=m+1}^{T} e_l y_{l-1} \right)^2 \right| \right)$$

$$= \left( O_p \left( \frac{1}{T} \right) + O_p \left( \frac{1}{M_T} \right) \right) (o_p(1) + o_p(1)) = O_p \left( \frac{1}{M_T} \right) = o_p(1).$$

Because $A_1 - A_6$ are $o_p(1)$ and $\lambda^2$ is a positive constant, we have

$$\Pr(|\hat{k} - k_0| > M_T) \leq \Pr(o_p(1) > \lambda^2) + \Pr(o_p(1) > \lambda^2) \to 2 \Pr(\lambda^2 < 0) = 0.$$ 

This means $\alpha = 0$, which contradicts the original argument that $\alpha > 0$. Thus $\hat{\tau}_T$ is $T$-consistent.

To find the limiting distributions of $\hat{\beta}_1(\hat{\tau}_T)$ and $\hat{\beta}_2(\hat{\tau}_T)$, note that $\hat{\tau}_T - \tau_0 = O_p(1/T)$ and

$$\sqrt{T} (\hat{\beta}_1(\hat{\tau}_T) - \hat{\beta}_1(\tau_0)) = \sqrt{T} \left( \frac{\sum_{l=1}^{[\hat{\tau}_T]} y_{l-1} y_{l-1}}{\sum_{l=1}^{[\hat{\tau}_T]} y_{l-1}^2} - \frac{\sum_{l=1}^{[\tau_0]} y_{l-1} y_{l-1}}{\sum_{l=1}^{[\tau_0]} y_{l-1}^2} \right)$$

$$= 1\{\hat{\tau}_T \leq \tau_0\} \sqrt{T} \left( \frac{\sum_{l=1}^{[\hat{\tau}_T]} e_l y_{l-1}}{\sum_{l=1}^{[\tau_0]} e_l y_{l-1}} - \frac{\sum_{l=1}^{[\tau_0]} e_l y_{l-1}}{\sum_{l=1}^{[\tau_0]} y_{l-1}^2} \right)$$

$$+ 1\{\hat{\tau}_T > \tau_0\} \sqrt{T} \left( \frac{\sum_{l=1}^{[\tau_0]} (\beta_1 y_{l-1} + e_l) y_{l-1}}{\sum_{l=1}^{[\tau_0]} y_{l-1}} + \sum_{l=[\tau_0]+1}^{[\hat{\tau}_T]} (\beta_2 y_{l-1} + e_l) y_{l-1} \right)$$

$$+ \frac{\sum_{l=1}^{[\tau_0]} (\beta_1 y_{l-1} + e_l) y_{l-1}}{\sum_{l=1}^{[\tau_0]} y_{l-1}} - \left( \sum_{l=1}^{[\tau_0]} y_{l-1}^2 \right)$$
Thus, $\beta_1(e_T)$ and $\beta_1(\tau_0)$ have the same asymptotic distribution. Similarly, $\beta_2(e_T)$ and $\beta_2(\tau_0)$ have the same asymptotic distribution.

Define $\mathcal{F}(t) = \sigma(e_i; i \leq t)$ as the sigma field generated by the past history of $\{e_i\}$. Because $\{e_i, y_{i-1}, \ldots, y_{i-T}\} \rightarrow \mathcal{F}$ and $\{e_i, y_{i-1}, \ldots, y_{i-T}\} \rightarrow \mathcal{F}$ are martingale difference sequences with

$$E(e_i, y_{i-1} | \mathcal{F}(t)) = 0 \quad (t = 1, 2, \ldots, T),$$

$$\frac{1}{[\tau_0 T]} \sum_{t=1}^{[\tau_0 T]} E((e_i, y_{i-1})^2 | \mathcal{F}(t)) \to \frac{\sigma^4}{1 - \beta_1^2} < \infty,$$

$$\frac{1}{T - [\tau_0 T]} \sum_{t=[\tau_0 T]+1}^{T} E((e_i, y_{i-1})^2 | \mathcal{F}(t)) \to \frac{\sigma^4}{1 - \beta_2^2} < \infty.$$

Applying the central limit theorem for martingale difference sequences (see, e.g., White, 1984, ch. V) and by (3b) and (3c) in Lemma 1, we have

$$\sqrt{T}(\hat{\beta}_1(\tau_T) - \beta_1) \stackrel{D}{=} \sqrt{T}(\hat{\beta}_1(\tau_0) - \beta_1) = \frac{\sum_{t} e_t y_{i-1}/\sqrt{T}}{\sum_{t} y_{i-1}/T} \to N \left(0, \frac{1 - \beta_1^2}{\tau_0} \right),$$

$$\sqrt{T}(\hat{\beta}_2(\tau_T) - \beta_2) \stackrel{D}{=} \sqrt{T}(\hat{\beta}_2(\tau_0) - \beta_1) = \frac{\sum_{t} e_t y_{i-1}/\sqrt{T}}{\sum_{t} y_{i-1}/T} \to N \left(0, \frac{1 - \beta_2^2}{\tau_0} \right).$$

This proves Theorem 1.
APPENDIX D: PROOF OF THEOREM 2

To derive the limiting distribution of \( \hat{\tau}_T \) for shrinking shift, let \( \beta_{2T} = \beta_1 + (1/\sqrt{g(T)}) \), where \( g(T) > 0 \), with \( g(T) \to \infty \) and \( [g(T)/T] \to 0 \) as \( T \to \infty \).

For \( \tau = \tau_0 + v[g(T)/T] \) and \( v \leq 0 \), by (B.2) in Appendix B and the facts that \( (i) \quad \lambda_T(\tau) = \left( \frac{\sum_{i=1}^{[\tau T]} y_{i-1} e_i}{\sum_{i=1}^{[\tau T]} y_{i-1}^2} \right)^2 \left( 1 - \frac{\sum_{i=1}^{[\tau T]} y_{i-1} e_i}{\sum_{i=1}^{[\tau T]} y_{i-1}^2} \right)^2 \right) \)

\[ + \left( \frac{\sum_{[\tau T]+1}^{T} e_i y_{i-1}}{\sum_{[\tau T]+1}^{T} y_{i-1}^2} \right)^2 \left( 1 - \frac{\sum_{[\tau T]+1}^{T} y_{i-1} e_i}{\sum_{[\tau T]+1}^{T} y_{i-1}^2} \right)^2 \] \[ = O_p(1)(1 - (1 - o_p(1))^2(1 + o_p(1)) \right) \]

\[ + O_p(1)(1 - (1 + o_p(1))^2(1 - o_p(1))) = o_p(1); \]

\[ (ii) \quad (\beta_{2T} - \beta_1) \left( \frac{\sum_{[\tau T]}^{[\tau_0 T]} y_{i-1}^2}{\sum_{[\tau T]}^{[\tau_0 T]} y_{i-1}^2} \right) \sum_{[\tau T]+1}^{T} y_{i-1} e_i \]

\[ = O_p(g(T))O_p(\sqrt{T}) \sqrt{g(T)}O_p(T) = O_p\left( \sqrt{g(T)} \right) \]

\[ = o_p(1); \]

\[ (iii) \quad \sum_{[\tau T]+1}^{T} y_{i-1}^2 \to 1, \]

we have

\[ RSS_T(\tau) - RSS_T(\tau_0) \]

\[ = -2(\beta_{2T} - \beta_1) \sum_{[\tau T]+1}^{[\tau_0 T]} y_{i-1} e_i + (\beta_{2T} - \beta_1)^2 \sum_{[\tau T]+1}^{[\tau_0 T]} y_{i-1}^2 + o_p(1) \]

\[ = -\frac{2}{\sqrt{g(T)}} \sum_{0}^{[\nu g(T)]-1} y_{k_0-t-1} e_{k_0-t} + \frac{1}{g(T)} \sum_{0}^{[\nu g(T)]-1} y_{k_0-t-1}^2 + o_p(1) \]

\[ \Rightarrow -2\sigma^2 \left( B_1 \left( \frac{|v|}{1 - \beta_1^2} \right) - \frac{1}{2} \frac{|v|}{1 - \beta_1^2} \right), \]

where \( B_1(\cdot) \) is a Brownian motion defined on \( R_+ \).
Similarly, for \( \tau = \tau_0 + \frac{v[g(T)/T]}{u} \) and \( u > 0 \), from (B.4) in Appendix B and the fact that \( [u/(1 - \beta_2^2)] \rightarrow [u/(1 - \beta_1^2)] \) as \( T \rightarrow \infty \), we have

\[
RSS_T(\tau) - RSS_T(\tau_0) = 2(\beta_{2T} - \beta_1) \sum_{[\tau T]} y_{t-1}e_t = (\beta_{2T} - \beta_1)^2 \sum_{[\tau T]} y_{t-1}^2 + o_p(1)
\]

\[
= -\frac{2}{\sqrt{g(T)}} \sum_{0}^{[v(x(T)]-1} (-y_{k_0+r}e_{k_0+r+1}) + \frac{1}{g(T)} \sum_{0}^{[v(x(T)]-1} y_{k_0+r}^2 + o_p(1)
\]

\[
= -2\sigma^2 \left(B_2 \left( \frac{v}{1 - \beta_1^2} \right) - \frac{v}{2(1 - \beta_1^2)} \right),
\]

where \( B_2(.) \) is another Brownian motion defined on \( R_+ \) independent of \( B_1(.) \).

Define \( r = u/(1 - \beta_1^2) \) and apply the continuous mapping theorem for argmax functionals (see Kim and Pollard, 1990). We have

\[
\frac{T(\beta_{2T} - \beta_1)^2}{1 - \beta_1^2} (\hat{\tau}_T - \tau_0) = \hat{r} = \text{Arg min}_r \{ RSS_T(\tau) - RSS_T(\tau_0) \}
\]

\[
d \rightarrow \text{Arg min}_r \left\{ -2\sigma^2 \left(B^*(r) - \frac{1}{2} |r| \right) \right\}
\]

\[
= \text{Arg max}_r \left( B^*(r) - \frac{1}{2} |r| \right),
\]

where \( B^*(r) \) is a two-sided Brownian motion on \( R \) defined to be \( B^*(r) = B_1(-r) \) for \( r \leq 0 \) and \( B^*(r) = B_2(r) \) for \( r > 0 \). This proves Theorem 2.

**APPENDIX E: PROOF OF LEMMA 2**

Conditions (4a) and (4d) are special cases of (3a). Conditions (4e) and (4f) are identical to (3d) and (3e), respectively.

\[
(4b) \quad \frac{1}{T} \sum_{[\tau T]} y_{t-1}e_t = \frac{1}{2T} \sum_{[\tau T]} (y_t^2 - y_{t-1}^2 - \varepsilon_t^2)
\]

\[
= \frac{T - k_0}{2T} \left( \frac{y_k^2}{T - k_0} - \frac{y_{k_0}^2}{T - k_0} + \frac{\sum_{[\tau T]} \varepsilon_t^2}{T - k_0} \right)
\]

\[
= \frac{1}{2} - \tau_0 \left( \frac{y_{k_0}^2}{T - k_0} + \frac{1}{\sqrt{T - k_0}} \sum_{j=k_0+1}^{T} \varepsilon_j \right)^2
\]

\[
- \frac{y_{k_0}^2}{T - k_0} - \frac{1}{T - k_0} \sum_{[\tau T]} \varepsilon_t^2.
\]
Because
\[
\frac{y_{k_0}}{\sqrt{T-k_0}} \overset{p}{\to} 0 \quad \text{and} \quad \frac{1}{\sqrt{T-k_0}} \sum_{j=k_0+1}^{T} e_j \Rightarrow \sigma B(1),
\]
by the continuous mapping theorem, we have
\[
\left( \frac{y_{k_0}}{\sqrt{T-k_0}} + \frac{1}{\sqrt{T-k_0}} \sum_{j=k_0+1}^{T} e_j \right)^2 \Rightarrow \sigma^2 B^2(1).
\]

Therefore
\[
\frac{1}{T} \sum_{[\tau_0 T]+1}^{T} y_{t-1} e_t \Rightarrow \frac{(1-\tau_0) \sigma^2}{2} (B^2(1) - 1) = O_p(1).
\]

(4c) \[
\frac{1}{T^2} \sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 = \frac{(T-k_0)^2}{T^2} \sum_{[\tau_0 T]+1}^{T} \left( \frac{y_{k_0}}{\sqrt{T-k_0}} + \frac{1}{\sqrt{T-k_0}} \sum_{j=k_0+1}^{T} e_j \right)^2 \frac{1}{T-k_0}
\]

\[
= (1-\tau_0)^2 \sum_{[\tau_0 T]+1}^{T} \left( o_p(1) + \frac{1}{\sqrt{T-k_0}} \sum_{j=0}^{r-k_0+1} e_{k_0+j+1} \right)^2 \frac{1}{T-k_0}
\]

\[
\Rightarrow (1-\tau_0)^2 \sigma^2 \int_0^1 B^2(s) \, ds = O_p(1),
\]

where \( s = (t - k_0 - 2)/(T - k_0) \).

To prove (4g), we use the triangle inequality, (4b)–(4e), and the facts that
\[
\sup_{[\tau_0 T]} \sum_{[\tau_0 T]} y_{t-1}^2 \leq \sum_{1} y_{t-1}^2.
\]

Then we have
\[
\sup_{\tau \leq \tau \leq \tau_0} |\hat{\beta}_2(\tau) - 1| = \sup_{\tau \leq \tau \leq \tau_0} \left| \frac{(\bar{\beta}_1 - 1) \sum_{[\tau_0 T]} y_{t-1}^2 + \sum_{[\tau_0 T]} \epsilon_t y_{t-1} + \sum_{[\tau_0 T]} \epsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} \right|
\]

\[
\leq \frac{1}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2} \left( \sum_{[\tau_0 T]} \bar{\beta}_1^2 - 1 \right) \sum_{1} y_{t-1}^2 + \sup_{\tau \leq \tau \leq \tau_0} \left| \sum_{[\tau_0 T]+1}^{T} \epsilon_t y_{t-1} \right|
\]

\[
= O_p(T) + o_p(T) + O_p(T) = O_p\left( \frac{1}{T} \right).
\]
(4h) By the triangle inequality and (4g),

\[
\sup_{\tau \in \Xi_T} |\hat{\beta}_2(\tau) - \beta_1| \leq \sup_{\tau \in \Xi_T} |\hat{\beta}_2(\tau) - 1| + |1 - \beta_1| = O_p(1).
\]

Conditions (4i) and (4j) can be proved by using an argument similar to that in (4b).

(4k) For \(\tau \in \Xi_T\),

\[
|\hat{\beta}_1(\tau) - 1| = \left| \frac{\sum_{1}^{[\tau_0 \tau]} y_{i-1}^2 + \sum_{1}^{[\tau_0 \tau]} \varepsilon_i y_{i-1} + \sum_{1}^{[\tau_0 \tau]} \varepsilon_i y_{i-1}}{\sum_{1}^{[\tau_0 \tau]} y_{i-1}^2 + \sum_{1}^{[\tau_0 \tau]} y_{i-1}^2} \right| \leq |\beta_1 - 1| \cdot \frac{1}{T} \sum_{1}^{[\tau_0 \tau]} y_{i-1}^2 \frac{T}{(\tau - \tau_0 T)^2} A_T + \frac{T}{[\tau_0 \tau]} \sum_{1}^{[\tau_0 \tau]} \varepsilon_i y_{i-1} \frac{T}{T}.
\]

where

\[
A_T = \frac{1}{\sum_{0}^{[\tau_0 \tau]}} \left( \frac{y_{k_0 + \tau}}{\sqrt{[(\tau - \tau_0 T)]}} \right)^2 \frac{1}{[\tau_0 \tau] T}.
\]

The proof is completed by utilizing (4a), (4c), (4i) and the following facts.

(i) By the definition of \(\Xi_T\), \(T/[(\tau - \tau_0 T)^2] = o(1)\) for \(\tau \in \Xi_T\).

(ii) By the definition of \(\Xi_T\) and by using the continuous mapping theorem, we have

\[
A_T = \frac{1}{\sigma^2 \int_{0}^{1} B^2} = O_p(1).
\]

(4l) By the fact that

\[
\sup_{\tau \in \Xi_T} \left| \frac{T^2}{\sum_{1}^{[\tau_0 \tau]} y_{i-1}^2} \right| \leq \left| \frac{T^2}{\sum_{1}^{[\tau_0 \tau]} y_{i-1}^2} \right| = O_p(1)
\]
and by (4j),

$$\sup_{\tau \in \mathbb{R}_+} |\hat{\beta}_2(\tau) - 1| = \sup_{\tau \in \mathbb{R}_+} \left| \frac{\sum_{[\tau T] + 1}^{T} y_{t-1} \varepsilon_t}{\sum_{k+1}^{T} y_{t-1}^2} \right| \leq \frac{1}{T} \left( \frac{T^2}{\sum_{k+1}^{T} y_{t-1}^2} \sup_{\tau \in \mathbb{R}_+} \left| \frac{\sum_{[\tau T] + 1}^{T} y_{t-1} \varepsilon_t}{T} \right| \right) = O_p \left( \frac{1}{T} \right).$$

These prove Lemma 2.

APPENDIX F:
DERIVATION OF EQUATIONS (12)–(14)

For $|\beta_1| < 1$, $\beta_2 = 1$, by (B.1) in Appendix B, the triangle inequality, and the fact that $\sup |ab| \leq \sup |a| \sup |b|$, we have

$$\sup_{\tau \leq \tau_0} \left| \frac{1}{T} \sum_{1}^{T} \varepsilon_t^2 - \sigma^2 \right| \leq \sup_{\tau \leq \tau_0} \left| \frac{1}{T} \sum_{1}^{T} \varepsilon_t^2 - \sigma^2 \right| + \sup_{\tau \leq \tau_0} \left| \hat{\beta}_1(\tau) - \beta_1 \right| \sup_{\tau \leq \tau_0} \left| \frac{1}{T} \sum_{[\tau T] + 1}^{T} y_{t-1} \varepsilon_t \right|$$

$$+ 2 \sup_{\tau \leq \tau_0} \left| \hat{\beta}_2(\tau) - \beta_1 \right| \sup_{\tau \leq \tau_0} \left| \frac{1}{T} \sum_{[\tau T] + 1}^{T} y_{t-1} \varepsilon_t \right|$$

$$+ \sup_{\tau \leq \tau_0} \left| (\hat{\beta}_2(\tau) - \beta_1)^2 \frac{1}{T} \sum_{[\tau T] + 1}^{T} y_{t-1}^2 - \frac{(1 - \beta_1)(\tau_0 - \tau) \sigma^2}{1 + \beta_1} \right|$$

$$+ 2 \sup_{\tau \leq \tau_0} \left| \hat{\beta}_2(\tau) - 1 \right| \left| \frac{\sum_{[\tau_0 T] + 1}^{T} y_{t-1} \varepsilon_t}{T} \right|$$

$$+ \frac{1}{T} \left( \sup_{\tau \leq \tau_0} \left| T(\hat{\beta}_2(\tau) - 1) \right| \right)^2 \frac{\sum_{[\tau_0 T] + 1}^{T} y_{t-1}^2}{T^2}$$

$$= o_p(1).$$
All six preceding terms above are $o_p(1)$ by Lemma 2. To show that the fourth term is also $o_p(1)$, we use the fact that $(1 - \beta_1)/(1 + \beta_1) = (1 - \beta_1)^2/(1 - \beta_1^2)$ and the triangle inequality that $\sup |a - b| = \sup |a - c + c - b| \leq \sup |a - c| + \sup |c - b|$, with

$$a = (\hat{\beta}_2(\tau) - \beta_1)^2 \frac{\sum_{[\tau T]} y_{t-1}^2}{T}, \quad b = \frac{(1 - \beta_1)(\tau_0 - \tau)\sigma^2}{1 + \beta_1}, \quad \text{and}$$

$$c = (\hat{\beta}_2(\tau) - \beta_1)^2 \frac{(\tau_0 - \tau)\sigma^2}{1 - \beta_1^2}.$$

Conditions (4e) and (4h) in Lemma 2 imply that $\sup |a - c|$ and $\sup |c - b|$ are both $o_p(1)$. This derives (12).

For $\tau \in \Xi$, by (B.3) in Appendix B and the triangle inequality, we have

$$\left| \frac{1}{T} \text{RSS}_T(\tau) - \sigma^2 - \frac{(1 - \beta_1)\tau_0 \sigma^2}{1 + \beta_1} \right|$$

$$\leq \left| \frac{1}{T} \sum_{1}^{T} e_t^2 - \sigma^2 \right| + \left| 2(\hat{\beta}_1(\tau) - \beta_1) \frac{1}{T} \sum_{[\tau T]} y_{t-1} e_t \right|$$

$$+ \left| (\hat{\beta}_1(\tau) - \beta_1)^2 \frac{\sum_{[\tau T]} y_{t-1}^2}{T} - \frac{(1 - \beta_1)\tau_0 \sigma^2}{1 + \beta_1} \right| + \left| 2(\hat{\beta}_1(\tau) - 1) \frac{1}{T} \sum_{[\tau T] + 1} y_{t-1} e_i \right|$$

$$+ (\hat{\beta}_1(\tau) - 1)^2 \frac{1}{T} \sum_{[\tau T]} y_{t-1}^2 + \left| (\hat{\beta}_2(\tau) - 1) \frac{1}{T} \sum_{[\tau T] + 1} y_{t-1} e_i \right|.$$

In the third term, we use the facts that $(1 - \beta_1)/(1 + \beta_1) = [(1 - \beta_1)^2]/(1 - \beta_1^2)$ and the triangle inequality that $|a - b| = |a - c + c - b| \leq |a - c| + |c - b|$, with

$$a = (\hat{\beta}_1(\tau) - \beta_1)^2 \frac{\sum_{1}^{T} y_{t-1}^2}{T}, \quad b = \frac{(1 - \beta_1)\tau_0 \sigma^2}{1 + \beta_1} \quad \text{and} \quad c = (\hat{\beta}_1(\tau) - \beta_1)^2 \frac{\tau_0 \sigma^2}{1 - \beta_1^2}.$$

Thus, the preceding equation is bounded by

$$\leq \left| \frac{1}{T} \sum_{1}^{T} e_t^2 - \sigma^2 \right| + 2 \left| \frac{1}{T} \sum_{[\tau T]} y_{t-1} e_t \right| |\hat{\beta}_1(\tau) - \beta_1|$$

$$+ (\hat{\beta}_1(\tau) - \beta_1)^2 \frac{1}{T} \sum_{[\tau T]} y_{t-1}^2 - \frac{\tau_0 \sigma^2}{1 - \beta_1^2}$$

$$+ \frac{\tau_0 \sigma^2}{1 - \beta_1^2} |\hat{\beta}_1(\tau) + 1 - 2\beta_1||\hat{\beta}_1(\tau) - 1| + 2 |\hat{\beta}_1(\tau) - 1| \left| \frac{1}{T} \sum_{[\tau T] + 1} y_{t-1} e_i \right|$$

$$+ \frac{[(\tau - \tau_0)T]^2}{T} (\hat{\beta}_1(\tau) - 1)^2 \frac{1}{[(\tau - \tau_0)T]^2 \sum_{[\tau T] + 1} y_{t-1}^2}$$

$$+ \left| \hat{\beta}_2(\tau) - 1 \right| \left| \frac{1}{T} \sum_{[\tau T] + 1} y_{t-1} e_i \right|.$$
All seven preceding terms above are \( o_p(1) \) by Lemma 2 and by the definition of \( \Xi^T \). This derives equation (13).

To derive (14), first note that for any positive constant \( c \),

\[
\hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) = \theta_T \left( \frac{1}{2}, c \right) \left( \beta_1 + \frac{\sum_{l=1}^{k_0} e_l y_{l-1}}{\sum_{l=1}^{k_0} y_{l-1}^2} \right) + \left( 1 - \theta_T \left( \frac{1}{2}, c \right) \right) \left( 1 + \frac{\sum_{k_0+1}^{[k_0+c \sqrt{T}]} e_l y_{l-1}}{\sum_{k_0+1}^{k_0+c \sqrt{T}} y_{l-1}^2} \right),
\]

where

\[
\theta_T \left( \frac{1}{2}, c \right) = \frac{\sum_{l=1}^{k_0} y_{l-1}^2}{\sum_{l=1}^{k_0} y_{l-1}^2 + \sum_{k_0+1}^{[k_0+c \sqrt{T}]} y_{l-1}^2}.
\]

By (3b) and the fact that \( y_{k_0}^2 / T \xrightarrow{P} 0 \), plus the invariance principle that

\[
\frac{1}{T} \sum_{k_0}^{[k_0+c \sqrt{T}]} y_{l-1}^2 = \frac{1}{T} \sum_{l=1}^{c} y_{l+k_0-1} \Rightarrow \sigma^2 \int_0^c B_3^2,
\]

where \( B_3(\cdot) \) is a Brownian motion defined on \( R_+ \), we have

(i) \( \theta_T \left( \frac{1}{2}, c \right) \Rightarrow \frac{\tau_0}{\tau_0 + (1 - \beta_1^2) \int_0^c B_3^2} \),

(ii) \( \hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) \Rightarrow \frac{\tau_0 \beta_1 + (1 - \beta_1^2) \int_0^c B_3^2}{\tau_0 + (1 - \beta_1^2) \int_0^c B_3^2} \).

Additionally, we have

(iii) \( \sup_{c \in R^{++}} \left| \hat{\beta}_2 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - 1 \right| = O_p(1); \)

(iv) \( \sup_{c \in R^{++}} \left| \frac{1}{T} \sum_{k_0+1}^{[k_0+c \sqrt{T}]} y_{l-1} e_l \right| = O_p(1); \)
\[
\left( v \right) \sup_{c \in R_{+}} \frac{1}{T} \left| \sum_{k_0+c \sqrt{T}+1}^{T} y_{t-1} \varepsilon_t \right| = O_p(1).
\]

From (B.3) in Appendix B, (4e) of Lemma 2, and (i)--(v) given immediately preceding, we have

\[
\frac{1}{T} \text{RSS}_T \left( \tau_0 + \frac{c}{\sqrt{T}} \right) = \frac{1}{T} \sum_{1}^{T} \varepsilon_t^2 - 2 \left( \hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - \beta_1 \right) \frac{1}{T} \sum_{1}^{k_0} y_{t-1} \varepsilon_t
\]

\[
+ \left( \hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - \beta_1 \right)^2 \frac{1}{T} \sum_{1}^{k_0} y_{t-1}^2
\]

\[
- 2 \left( \hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - 1 \right) \frac{1}{T} \sum_{k_0+1}^{[k_0+c \sqrt{T}]} y_{t-1} \varepsilon_t
\]

\[
+ \left( \hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - 1 \right)^2 \frac{1}{T} \sum_{k_0+1}^{[k_0+c \sqrt{T}]} y_{t-1}^2
\]

\[
- \left( \hat{\beta}_2 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - 1 \right) \frac{1}{T} \sum_{[k_0+c \sqrt{T}]+1}^{T} y_{t-1} \varepsilon_t
\]

\[
= \frac{1}{T} \sum_{1}^{T} \varepsilon_t^2 + o_p(1) + \left( \hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - \beta_1 \right)^2 \frac{1}{T} \sum_{1}^{k_0} y_{t-1}^2 + o_p(1)
\]

\[
+ \left( \hat{\beta}_1 \left( \tau_0 + \frac{c}{\sqrt{T}} \right) - 1 \right)^2 \frac{1}{T} \sum_{k_0+1}^{[k_0+c \sqrt{T}]} y_{t-1}^2 + o_p(1)
\]

\[
\Rightarrow \sigma^2 + \left( \frac{(1-\beta_1)(1-\beta_1^2)}{\tau_0 + (1-\beta_1^2)} \int_{0}^{c} B_2^2 \right)^2 \frac{\tau_0 \sigma^2}{1-\beta_1^2}
\]

\[
+ \left( \frac{\tau_0(\beta_1-1)}{\tau_0 + (1-\beta_1^2)} \int_{0}^{c} B_3^2 \right)^2 \sigma^2 \int_{0}^{c} B_3^2
\]

\[
= \sigma^2 + \frac{\tau_0 \sigma^2 (1-\beta_1)^2}{\tau_0 \left( \int_{0}^{c} B_3^2 \right)^{-1} + 1-\beta_1^2}.
\]

This derives equation (14).
APPENDIX G:
PROOF OF THEOREM 3

To show $|\hat{\tau}_T - \tau_0| = O_p(1/T)$, it is sufficient to prove that $A_1 - A_6$ defined in Appendix C are all $o_p(1)$ in the case where $|\beta_1| < 1$ and $\beta_2 = 1$. Now

$$A_1 = \frac{\sum_{k_0+1}^{T} y_{t-1} \epsilon_t}{\sum_{k_0+1}^{T} y_{t-1}^2} = o_p(1)$$

by (4b) and (4c) of Lemma 2;

$$A_2 = \sup_{m \in D_{1T}} \frac{\sum_{m+1}^{k_0} y_{t-1} \epsilon_t}{\sum_{m+1}^{k_0} y_{t-1}^2} = o_p(1)$$

by the uniform law of large numbers in Andrews (1987, Theorem 1);

$$A_3 = \sup_{m \in D_{1T}} \left| \frac{\sum_{m+1}^{T} y_{t-1}^2}{\sum_{k_0+1}^{T} y_{t-1}^2} \Lambda_T \left( \frac{m}{T} \right) \right| \leq \left( \frac{1}{\sum_{k_0+1}^{T} y_{t-1}^2} + \frac{1}{\sum_{k_0-M_T}^{k_0} y_{t-1}^2} \right)$$

$$\times \left( \sup_{m \in D_{1T}} \left| \frac{\left( \sum_{1}^{k_0} \epsilon_{t-1} \right)^2}{\sum_{1}^{k_0} y_{t-1}^2} - \frac{\left( \sum_{1}^{m} \epsilon_{t-1} \right)^2}{\sum_{1}^{m} y_{t-1}^2} \right| \right)$$

$$+ \sup_{m \in D_{1T}} \left| \frac{\left( \sum_{k_0+1}^{T} \epsilon_{t-1} y_{t-1} \right)^2}{\sum_{k_0+1}^{T} y_{t-1}^2} - \frac{\left( \sum_{m+1}^{T} \epsilon_{t-1} y_{t-1} \right)^2}{\sum_{m+1}^{T} y_{t-1}^2} \right|$$

$$= \left( O_p \left( \frac{1}{T^2} \right) + O_p \left( \frac{1}{MT} \right) \right) (o_p(1) + O_p(1)) = O_p \left( \frac{1}{MT} \right) = o_p(1);$$
by (4a) and (4e) of Lemma 2;

$$A_5 = \sup_{m \in D_T} \left( \sum_{k_0+1}^{m} \sum_{1}^{k_0+1} \left( \frac{\sum_{1}^{k_0+1} y_{t-1} \varepsilon_t}{\sum_{1}^{k_0+1} y_{t-1}^2} - \sum_{m+1}^{T} \sum_{1}^{k_0+1} \frac{\varepsilon_t y_{t-1}}{\sum_{1}^{k_0+1} y_{t-1}^2} \right)^2 \right) + \sup_{m \in D_T} \left( \frac{\sum_{k_0+1}^{m} \sum_{1}^{k_0+1} \left( \frac{\sum_{1}^{k_0+1} y_{t-1} \varepsilon_t}{\sum_{1}^{k_0+1} y_{t-1}^2} - \sum_{m+1}^{T} \sum_{1}^{k_0+1} \frac{\varepsilon_t y_{t-1}}{\sum_{1}^{k_0+1} y_{t-1}^2} \right)^2}{\sum_{1}^{k_0+1} y_{t-1}^2} \right)$$

$$= \left( O_p \left( \frac{1}{T} \right) + O_p \left( \frac{1}{M_T^2} \right) \right) \left( O_p(1) + O_p(1) \right) = o_p(1).$$

(To prove that $A_5 = o_p(1)$, explicit formulae would be needed. I am not able to provide a general proof because of the unknown asymptotic properties of $\sup_{m \in D_T} \left( \sum_{k_0+1}^{m} \sum_{1}^{k_0+1} \left( \frac{\sum_{1}^{k_0+1} y_{t-1} \varepsilon_t}{\sum_{1}^{k_0+1} y_{t-1}^2} - \sum_{m+1}^{T} \sum_{1}^{k_0+1} \frac{\varepsilon_t y_{t-1}}{\sum_{1}^{k_0+1} y_{t-1}^2} \right)^2 \right)$ under nonstationarity of $y_t$. No study on the uniform convergence under this kind of nonstationarity has been done in the literature.)

Because $A_1-A_6$ are $o_p(1)$, thus $\hat{\tau}_T$ is $T$-consistent.

To find the limiting distribution of $\hat{\beta}_1(\hat{\tau}_T)$, note that $\hat{\tau}_T - \tau_0 = O_p(1/T)$ and

$$\sqrt{T} \left( \hat{\beta}_1(\hat{\tau}_T) - \hat{\beta}_1(\tau_0) \right)$$

$$= \sqrt{T} \left( \begin{bmatrix} \sum_{1}^{[\hat{\tau}_T]} y_t y_{t-1} \sum_{1}^{[\tau_0 T]} y_t y_{t-1} \\ \sum_{1}^{[\hat{\tau}_T]} y_t^2 \sum_{1}^{[\tau_0 T]} y_t^2 \end{bmatrix} \right).$$
Thus, \( \pi_t(\tau_0) \) and \( \pi_1(\tau_0) \) have the same asymptotic distribution. Because \( \{ct, yt-3\} \) is a martingale difference sequence, with

\[
E(\varepsilon_t, y_{t-1}|\mathcal{F}_{t-1}) = 0, \\
\sum_{t=1}^{[\tau_0 T]} E((\varepsilon_t, y_{t-1})^2|\mathcal{F}_{t-1}) \overset{p}{\to} \frac{\sigma^4}{1-\beta_1^2} < \infty.
\]

Applying the central limit theorem for martingale difference sequences and the fact that

\[
\frac{1}{T} \sum_{t=1}^{[\tau_0 T]} y_{t-1}^2 \overset{p}{\to} \tau_0 \sigma^2, \\
\sqrt{T}(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \overset{d}{=} \sqrt{T}(\hat{\beta}_1(\tau_0) - \beta_1) = \frac{\sum_{t=1}^{[\tau_0 T]} \varepsilon_t, y_{t-1}/\sqrt{T}}{\sum_{t=1}^{[\tau_0 T]} y_{t-1}^2/T} \overset{d}{\to} N(0, \frac{1-\beta_1^2}{\tau_0^2}).
\]
To find the limiting distribution of $\hat{\beta}_2(\hat{\tau}_T)$, note that $\hat{\tau}_T - \tau_0 = O_p(1/T)$ and

$$T(\hat{\beta}_2(\hat{\tau}_T) - \hat{\beta}_2(\tau_0))$$

$$= T \left( \frac{\sum_{[\tau_0 T] + 1}^{[\tau_T] + 1} y_{t} y_{t-1}}{\sum_{[\tau_T] + 1}^{[\tau_T] + 1} y_{t-1}^2} - \frac{\sum_{[\tau_0 T] + 1}^{[\tau_T] + 1} y_{t} y_{t-1}}{\sum_{[\tau_0 T] + 1}^{[\tau_0 T] + 1} y_{t-1}^2} \right)$$

$$= 1\{\hat{\tau}_T \leq \tau_0\} T \left( \frac{\sum_{[\tau_0 T] + 1}^{[\tau_T] + 1} y_{t} y_{t-1} - \sum_{[\tau_0 T] + 1}^{[\tau_T] + 1} y_{t} y_{t-1}}{\sum_{[\tau_T] + 1}^{[\tau_T] + 1} y_{t-1}^2} \right) + \frac{(\beta_1 - \beta_2) T}{\sum_{[\tau_T] + 1}^{[\tau_T] + 1} y_{t-1}^2}$$

$$+ 1\{\hat{\tau}_T > \tau_0\} T \left( \frac{\sum_{[\tau_0 T] + 1}^{[\tau_T] + 1} y_{t} y_{t-1} - \sum_{[\tau_0 T] + 1}^{[\tau_T] + 1} y_{t} y_{t-1}}{\sum_{[\tau_T] + 1}^{[\tau_T] + 1} y_{t-1}^2} \right)$$

$$= 1\{\hat{\tau}_T \leq \tau_0\} T \left( O_p \left( \frac{1}{T^2} \right) O_p \left( \frac{1}{T} \right) + O_p \left( \frac{1}{T^2} \right) \right)$$

$$+ 1\{\hat{\tau}_T > \tau_0\} T \left( O_p \left( \frac{1}{T^2} \right) O_p \left( \frac{1}{T} \right) - O_p \left( \frac{1}{T^2} \right) \right)$$

$$= o_p(1).$$

Thus, $\hat{\beta}_2(\hat{\tau}_T)$ and $\hat{\beta}_2(\tau_0)$ have the same asymptotic distribution. Applying (4b) and (4c) in Lemma 2, we get

$$T(\hat{\beta}_2(\hat{\tau}_T) - 1) \overset{d}{=} T(\hat{\beta}_2(\tau_0) - \beta_2) = \frac{1}{T} \sum_{[\tau_0 T] + 1}^{[\tau_T] + 1} \frac{\epsilon_t y_{t-1}}{y_{t-1}^2} - \frac{B^2(1) - 1}{2(1 - \tau_0) \int_0^1 B^2}.$$ 

These derive the limiting distribution of $\hat{\beta}_1(\hat{\tau}_T)$ and $\hat{\beta}_2(\hat{\tau}_T)$ for fixed magnitude of break.

To derive the limiting distribution of $\hat{\tau}_T$ for shrinking shift, we fix $\beta_2$ at one and let $\beta_{1T} = 1 - [1/g(T)]$, where $g(T) > 0$, with $g(T) \to \infty$ and $[g(T)/T] \to 0$ as $T \to \infty$. Let $\nu$ be a finite constant, and $B_1(\cdot)$ and $B_2(\cdot)$ be defined as in Section 2. For $\tau = \tau_0 + \nu[g(T)/T]$ and $\nu \leq 0$, we have $\Lambda_T(\tau) = o_p(1)$ where $\Lambda_T(\tau)$ is defined in Appendix B, and
\[ \frac{y_{k_0-t-1}}{\sqrt{g(T)}} = \frac{1}{\sqrt{g(T)}} \left[ \left( \frac{1}{1 - \frac{1}{g(T)}} \right)^{k_0-t-2} y_0 + \sum_{i=0}^{k_0-t-2} \left( \frac{1}{1 - \frac{1}{g(T)}} \right)^i g_{k_0-t-1-i} \right] \]

\[ = \sum_{i=0}^{k_0-t-2} \left( \left( \frac{1}{1 - \frac{1}{g(T)}} \right)^{g(T)} \frac{1}{\sqrt{g(T)}} e_{k_0-t-1-i} \right) + o_p(1) \]

\[ \Rightarrow \sigma \int_0^\infty \exp(-s) dB_1(s) \overset{d}{=} \sigma B_a \left( \frac{1}{2} \right), \]

where \( B_a \left( \frac{1}{2} \right) \) is a Brownian motion defined on \( \mathbb{R}_+ \).

Because

\[ \frac{y_{k_0-t-1}}{\sqrt{g(T)}} \Rightarrow \sigma B_a \left( \frac{1}{2} \right) \text{ and } \frac{1}{\sqrt{g(T)}} \sum_{0}^{[\lceil v/g(T) \rceil - 1]} e_{k_0-t} \Rightarrow \sigma B_1 (|v|), \]

we have

\[ \frac{1}{\sqrt{g(T)}} \sum_{0}^{[\lceil v/g(T) \rceil - 1]} \frac{y_{k_0-t-1} e_{k_0-t}}{\sqrt{g(T)}} \Rightarrow \sigma^2 B_a \left( \frac{1}{2} \right) B_1 (|v|) \text{ and } \]

\[ \frac{1}{g^2(T)} \sum_{0}^{[\lceil v/g(T) \rceil - 1]} y_{k_0-t-1}^2 \Rightarrow |v| \sigma^2 B_a^2 \left( \frac{1}{2} \right). \]

Further, because

\[ (1 - \beta_{1T}) \sum_{[\tau T] + 1}^{[\tau T]} \frac{y_{t-1}^2}{[\tau T] + 1} \sum_{[\tau T] + 1}^{T} y_{t-1} e_t = \frac{O_p(g^2(T)) O_p(T)}{g(T) O_p(T^2)} = O_p \left( \frac{g(T)}{T} \right) = o_p(1) \]

and

\[ \sum_{[\tau T] + 1}^{T} \frac{y_{t-1}^2}{[\tau T] + 1} \rightarrow 1, \]

equation (B.2) in Appendix B becomes

\[ RSS_T(\tau) - RSS_T(\tau_0) \]

\[ = -2(1 - \beta_{1T}) \sum_{[\tau T] + 1}^{[\tau T]} y_{t-1} e_t + (1 - \beta_{1T})^2 \sum_{[\tau T] + 1}^{[\tau T]} y_{t-1}^2 + o_p(1) \]

\[ = -2 \frac{[\lceil v/g(T) \rceil - 1]}{g(T)} \sum_{0}^{[\lceil v/g(T) \rceil - 1]} y_{k_0-t-1} e_{k_0-t} + \frac{1}{g^2(T)} \sum_{0}^{[\lceil v/g(T) \rceil - 1]} y_{k_0-t-1}^2 + o_p(1) \]

\[ \Rightarrow -2\sigma^2 B_a \left( \frac{1}{2} \right) B_1 (|v|) + |v| \sigma^2 B_a^2 \left( \frac{1}{2} \right). \]
Similarly, for $\tau = \tau_0 + \nu[g(T)/T]$ and $\nu > 0$, from (B.4) in Appendix B, we have

$$\Lambda_\tau(\tau) = o_p(1),$$

$$(1 - \beta_{1T}) \sum_{[\tau_0+1]}^{[\tau]} y_{t-1} e_t + \sum_{[\tau]}^{[\tau_0+1]} y_{t-1}^2 = O_p(T)O_p(g^2(T)) \frac{g(T)}{O_p(T^2)} = O_p \left( \frac{g(T)}{T} \right) = o_p(1),$$

$$\frac{y_{k_0}}{\sqrt{g(T)}} \left( 1 - \frac{1}{g(T)} \right)^{k_0} \frac{y_0}{\sqrt{g(T)}} + \sum_{j=0}^{k_0-1} \left( 1 - \frac{1}{g(T)} \right)^j \frac{\varepsilon_{k_0-i}}{\sqrt{g(T)}}$$

$$\Rightarrow \sigma^2 \int_0^\infty \exp(-s) dB_1(s) \frac{d}{ds} \sigma B_a \left( \frac{1}{2} \right),$$

$$\frac{1}{g(T)} \sum_{0}^{[vg(T)]-1} \left( \frac{y_{k_0+t}}{\sqrt{g(T)}} \right)^2 = \frac{1}{g(T)} \sum_{0}^{[vg(T)]-1} \left( \frac{1}{\sqrt{g(T)}} \sum_{i=0}^{t-1} \varepsilon_{k_0+t-i} + \frac{y_{k_0}}{\sqrt{g(T)}} \right)^2$$

$$\Rightarrow \sigma^2 \int_0^\nu \left( B_2(s) + B_a \left( \frac{1}{2} \right) \right)^2 ds,$$

$$\frac{1}{\sqrt{g(T)}} \sum_{0}^{[vg(T)]-1} \frac{y_{k_0+t}}{\sqrt{g(T)}} \varepsilon_{k_0+t+1} \Rightarrow \sigma^2 \int_0^\nu \left( B_2(s) + B_a \left( \frac{1}{2} \right) \right) dB_2(s).$$

Hence, equation (B.4) in Appendix B becomes

$$RSS_\tau(\tau) - RSS_\tau(\tau_0)$$

$$= 2(1 - \beta_{1T}) \sum_{[\tau_0+1]}^{[\tau]} y_{t-1} e_t + (1 - \beta_{1T})^2 \sum_{[\tau_0+1]}^{[\tau]} y_{t-1}^2 + o_p(1)$$

$$= 2 \frac{[vg(T)]-1}{g(T)} \sum_{0}^{[vg(T)-1]} y_{k_0+t} \varepsilon_{k_0+t+1} + \frac{1}{g^2(T)} \sum_{0}^{[vg(T)]-1} y_{k_0+t}^2 + o_p(1)$$

$$\Rightarrow 2\sigma^2 \int_0^\nu \left( B_2(s) + B_a \left( \frac{1}{2} \right) \right) dB_2(s) + \sigma^2 \int_0^\nu \left( B_2(s) + B \left( \frac{1}{2} \right) \right)^2 ds$$

$$= -2\sigma^2 B_a^2 \left( \frac{1}{2} \right)$$

$$\times \left[ -\frac{B_2(\nu)}{B_a \left( \frac{1}{2} \right)} - \int_0^\nu \frac{B_2(s)}{B_a^2 \left( \frac{1}{2} \right)} dB_2(s) - \int_0^\nu \left( \frac{B_2(s)}{2B_a \left( \frac{1}{2} \right)} + 1 \right) \frac{B_2(s)}{B_a \left( \frac{1}{2} \right)} ds - \frac{1}{2} \nu \right].$$
By the continuous mapping theorem for argmax functionals, we have

\[(1 - \beta_{1T})T(\hat{\tau}_T - \tau_0) = \hat{\nu} = \operatorname{Arg max}\{\text{RSS}_T(\tau) - \text{RSS}_T(\tau_0)\}\]

\[
\overset{d}{\to} \operatorname{Arg max}\left\{ -2\sigma^2B_\nu^2\left(\frac{1}{2}\right)\left(\frac{C^*(\nu)}{B_0\left(\frac{1}{2}\right)} - \frac{1}{2}|\nu|\right) \right\}
\]

\[= \operatorname{Arg max}\left(\frac{C^*(\nu)}{B_0\left(\frac{1}{2}\right)} - \frac{1}{2}|\nu|\right).\]

This proves Theorem 3.

\[\blacksquare\]

**APPENDIX H:**

**PROOF OF LEMMA 3**

(5a) By the fact that \((y_{k_0}\sqrt{T}) \Rightarrow \sigma B(\tau_0)\) and \((y_{0}^{2}/T) = o_p(1)\), we have

\[
\frac{1}{T} \sum_{1}^{\lfloor \tau_0 T \rfloor} y_{t-1} \epsilon_t = \frac{1}{2T} \sum_{1}^{\lfloor \tau_0 T \rfloor} (y_{t}^{2} - y_{t-1}^{2} - \epsilon_t^{2}) = \frac{1}{2T} \left(y_{k_0}^{2} - y_{0}^{2} - \sum_{1}^{\lfloor \tau_0 T \rfloor} \epsilon_t^{2}\right)
\]

\[
\Rightarrow \frac{\sigma^2}{2} (B^2(\tau_0) - \tau_0) = O_p(1).
\]

(5b) \[
\frac{1}{\sqrt{T}} \sum_{\lfloor \tau_0 T \rfloor + 1}^{T} y_{t-1} \epsilon_t
\]

\[
= \frac{1}{\sqrt{T}} \sum_{t = k_0 + 1}^{T} \left(\beta_2^{t-k_0-1} y_{k_0} + \sum_{i = k_0 + 1}^{t-1} \beta_2^{t-i-1} \epsilon_i\right) \epsilon_t
\]

\[
= \frac{1}{\sqrt{T}} \left(\sum_{j = 1}^{k_0} \left(\epsilon_j \sum_{k_0 + 1}^{T} \beta_2^{t-k_0-1} \epsilon_t\right) + \sum_{t = k_0 + 1}^{T} \sum_{i = k_0 + 1}^{t-1} \beta_2^{t-i-1} \epsilon_i \epsilon_t\right)
\]

\[
= \frac{1}{\sqrt{T}} \sum_{j = 1}^{k_0} \left(\epsilon_j \sum_{k_0 + 1}^{T} \beta_2^{t-k_0-1} \epsilon_t\right) + \frac{1}{\sqrt{T}} \sum_{t = k_0 + 1}^{T} \left(\epsilon_t \sum_{i = k_0 + 1}^{t-1} \beta_2^{t-i-1} \epsilon_i\right) + o_p(1).
\]
By Assumption (A2) that $\varepsilon$'s are i.i.d., \( \{ \varepsilon_j \sum_{t=k_0+1}^T \beta_{2}^{t-k_0-1} \varepsilon_t, \mathcal{F}_j \}_{j=1}^{k_0} \) and \( \{ \varepsilon_j \sum_{t=k_0+1}^T \beta_{2}^{t-k_0-1} \varepsilon_t, \mathcal{F}_j \}_{j=1}^{k_0} \) will both be martingale difference sequences with

\[
E \left( \varepsilon_j \sum_{t=k_0+1}^T \beta_{2}^{t-k_0-1} \varepsilon_t | \mathcal{F}_{j-1} \right) = 0 \quad (j = 1, 2, \ldots, k_0),
\]

\[
\frac{1}{T} \sum_{j=1}^{k_0} E \left( \left( \varepsilon_j \sum_{t=k_0+1}^T \beta_{2}^{t-k_0-1} \varepsilon_t \right)^2 | \mathcal{F}_{j-1} \right) \to \frac{\tau_0 \sigma^4}{1 - \beta_2^2} < \infty,
\]

\[
E \left( \varepsilon_j \sum_{t=k_0+1}^T \beta_{2}^{t-k_0-1} \varepsilon_t | \mathcal{F}_{j-1} \right) = 0 \quad (t = k_0 + 1, k_0 + 2, \ldots, T),
\]

\[
\frac{1}{T} \sum_{t=k_0+1}^T E \left( \left( \varepsilon_t \sum_{i=k_0+1}^{t-1} \beta_{2}^{i-k_0-1} \varepsilon_i \right)^2 | \mathcal{F}_{t-1} \right) \to \frac{(1 - \tau_0) \sigma^4}{1 - \beta_2^2}.
\]

By the central limit theorem for martingale difference sequences and by the independence of the two martingale difference sequences given previously, we have

\[
\frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t-1} \varepsilon_t \overset{d}{\to} N \left( 0, \frac{\sigma^4}{1 - \beta_2^2} \right);
\]

\[
(5c) \quad \frac{1}{T^2} \sum_{t=1}^{[r_0 T]} y_{t-1}^2 = \sum_{t=1}^{[r_0 T]} \left( \frac{\sum_{i=1}^{t-1} \varepsilon_i}{\sqrt{T}} + \frac{\varepsilon_0}{\sqrt{T}} \right)^2 \frac{1}{T} \to \alpha^2 \int_0^{r_0} B^2 = O_p(1);
\]

\[
(5d) \quad \frac{1}{T} \sum_{t=k_0+1}^{[r T]} y_{t-1}^2 = \frac{1}{T} \sum_{t=k_0+1}^{[r T]} \left( \beta_{2}^{t-k_0-1} y_{k_0} + \sum_{i=k_0+1}^{t-1} \beta_{2}^{i-k_0-1} \varepsilon_i \right)^2
\]

\[
= \left( \frac{y_{k_0}}{\sqrt{T}} \right)^2 \sum_{t=k_0+1}^{[r T]} \beta_{2}^{2(t-k_0-1)}
\]

\[
+ 2 \frac{y_{k_0}}{T} \sum_{t=k_0+1}^{[r T]} \left( \beta_{2}^{t-k_0-1} \sum_{i=k_0+1}^{t-1} \beta_{2}^{i-k_0-1} \varepsilon_i \right)
\]

\[
+ \frac{1}{T} \sum_{t=k_0+1}^{[r T]} \left( \beta_{2}^{t-k_0-1} \sum_{i=k_0+1}^{t-1} \beta_{2}^{i-k_0-1} \varepsilon_i \right)^2.
\]

The first term converges weakly to \( \left[ \sigma^2 B^2(\tau_0) \right] / (1 - \beta_2^2) \).

Because \( |y_{k_0} / \sqrt{T}| = O_p(1) \) and \( (1/\sqrt{T}) \sup_{r > k_0} |\varepsilon_r| = o_p(1) \), the second term is bounded by

\[
\leq 2 \frac{y_{k_0}}{\sqrt{T}} \sup_{r \in \mathbb{Z}^+} \left( \sum_{t=k_0+1}^{[r T]} \beta_{2}^{t-k_0-1} \frac{1 - \beta_{2}^{t-k_0-1}}{1 - \beta_2} \right) \frac{1}{\sqrt{T}} \sup_{r > k_0} |\varepsilon_r|,
\]

\[
\leq 2 \frac{y_{k_0}}{\sqrt{T}} \left( \left( \sum_{t=k_0+1}^{[r T]} \left| \beta_{2}^{t-k_0-1} \right| \frac{1}{1 - \beta_2} \right) + \left( \sum_{t=k_0+1}^{[r T]} \left| \beta_{2}^{2(t-k_0-1)} \right| \frac{1}{1 - \beta_2} \right) \right) \frac{1}{\sqrt{T}} \sup_{r > k_0} |\varepsilon_r|,
\]

\[
\leq 2 \frac{y_{k_0}}{\sqrt{T}} \left( \frac{1}{(1 - |\beta_2|)^2} + \frac{1}{(1 - |\beta_2|)(1 - \beta_2^2)} \right) \frac{1}{\sqrt{T}} \sup_{r > k_0} |\varepsilon_r| = o_p(1).
\]
Last, because $|\beta_2| < 1$ and $\varepsilon_i$ are i.i.d. with finite fourth moment, we have

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=k_0+1}^{[rT]} E \left( \sum_{i=k_0+1}^{t-1} \beta_2^{t-i} \varepsilon_i \right)^2 = \frac{(\tau - \tau_0)\sigma^2}{1 - \beta_2^2},$$

$$\text{Var} \left( \left( \sum_{i=k_0+1}^{t-1} \beta_2^{t-i} \varepsilon_i \right)^2 \right) < E \left( \sum_{i=k_0+1}^{t-1} \beta_2^{t-i} \varepsilon_i \right)^4 < \infty.$$ 

By the uniform weak law of large numbers for dependent and heterogeneous processes,

$$\sup_{\tau \in \mathbb{Z}_+} \left| \frac{1}{T} \sum_{t=k_0+1}^{[rT]} \left( \sum_{i=k_0+1}^{t-1} \beta_2^{t-i} \varepsilon_i \right)^2 - \frac{(\tau - \tau_0)\sigma^2}{1 - \beta_2^2} \right| = o_p(1).$$

Thus,

$$\frac{1}{T} \sum_{t=k_0+1}^{[rT]} y_{t-1}^2 \Rightarrow \frac{\tau - \tau_0 + B^2(\tau_0)}{1 - \beta_2^2} \sigma^2.$$

(5e) $\sup_{\tau \in \mathbb{Z}_+} \left| \frac{1}{T} \sum_{t=1}^{[\tau T]} y_{t-1} \varepsilon_t \right| = \sup_{\tau \in \mathbb{Z}_+} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{[\tau T]} \frac{y_{t-1} \varepsilon_t}{\sqrt{T}} \right| \overset{d}{\Rightarrow} \sup_{\tau \in \mathbb{Z}_+} \sigma^2 \int_0^\tau dB \bigg| = O_p(1);$

(5f) $\sup_{\tau \in \mathbb{Z}_+} \left| \frac{1}{T} \sum_{t=1}^{[\tau T]} y_{t-1} \varepsilon_t \right| = \sup_{\tau \in \mathbb{Z}_+} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{[\tau T]} \frac{y_{t-1} \varepsilon_t}{\sqrt{T}} \right| \overset{d}{\Rightarrow} \sup_{\tau \in \mathbb{Z}_+} \sigma^2 \int_0^{\tau_0} dB \bigg| = O_p(1);$

(5g) $\sup_{\tau \in \mathbb{Z}_+} |\hat{\beta}_1(\tau) - 1| = \sup_{\tau \in \mathbb{Z}_+} \left| \frac{1}{T} \sum_{t=1}^{[\tau T]} y_{t-1} \varepsilon_t \right| \leq \frac{1}{T} \left( \sum_{t=1}^{[\tau T]} y_{t-1}^2 \right)^{1/2} \sup_{\tau \in \mathbb{Z}_+} \left| \frac{1}{T} \sum_{t=1}^{[\tau T]} y_{t-1} \varepsilon_t \right|$

$$= \frac{1}{T} O_p(1) = O_p(1);$$

(5h) $\sup_{\tau \in \mathbb{Z}_+} |\hat{\beta}_2(\tau) - 1| = \sup_{\tau \in \mathbb{Z}_+} \left| \beta_2 - 1 \right| \left( \sum_{t=1}^{[\tau T]} y_{t-1}^2 + \sum_{t=1}^{[\tau T] + 1} \varepsilon_t y_{t-1} - \sum_{t=1}^{[\tau T]} y_{t-1} \right) \bigg|$

$$\leq \sup_{\tau \in \mathbb{Z}_+} \frac{\left| \sum_{t=1}^{[\tau T] + 1} \varepsilon_t y_{t-1} \right|}{\sum_{t=1}^{[\tau T] + 1} y_{t-1}} + \sup_{\tau \in \mathbb{Z}_+} \frac{\sum_{t=1}^{[\tau T]} \varepsilon_t y_{t-1}}{\sum_{t=1}^{[\tau T]} y_{t-1}^2} \bigg| + \left| \sum_{t=1}^{[\tau T] + 1} \varepsilon_t y_{t-1} \right| \bigg| \bigg| \bigg| \bigg|.
Using the fact that

\[
\sum_{[\tau T]+1}^{[\tau_0 T]} \left( \frac{y_{\tau-1}}{\sqrt{T}} \right)^2 \frac{1}{[(\tau_0 - \tau)T]} \Rightarrow \sigma^2 \int_\tau^{\tau_0} B^2 = O_p(1),
\]

the definition of \( \Xi^T \), (5d), and (5f), the first term is bounded by

\[
\sup_{\tau \in \Xi^T} \frac{|\beta_2 - 1|}{k_0 - k} \sup_{\tau \in \Xi^T} \frac{T(k_0 - k)}{[\tau_0 T]} \sum_{[\tau T]+1}^{[\tau_0 T]} y_{\tau-1}^2 = O_p(T^{-1/2})O_p(1)O_p(1) = o_p(1).
\]

The second term is bounded by

\[
= O_p(T^{-1/2})O_p(1)O_p(1) = o_p(1).
\]

The last term is in \( o_p(1) \) by (5b) and (5d) of this lemma. Thus, \( \sup_{\tau \in \Xi^T} |\hat{\beta}_2(\tau) - 1| = O_p(1) \).

(5i) \( \sup_{\tau \in \Xi^T} \left| \frac{1}{T} \sum_{[\tau T]+1}^{[\tau_0 T]} y_{\tau-1} e_t \right| = \sup_{\tau \in \Xi^T} \left| \frac{1}{T} \sum_{[\tau T]+1}^{[\tau_0 T]} \left( \beta_2^{T-k_0} y_{k_0} + \sum_{i=k_0+1}^{i-1} \beta_2^{i-1} e_i \right) e_t \right| \leq \sup_{\tau \in \Xi^T} \left| \frac{1}{T} \sum_{[\tau T]+1}^{[\tau_0 T]} \beta_2^{T-k_0} y_{k_0} e_t \right| + \sup_{\tau \in \Xi^T} \left| \frac{1}{T} \sum_{[\tau T]+1}^{[\tau_0 T]} \sum_{i=k_0+1}^{i-1} \beta_2^{i-1} e_i e_t \right|.
\]

Because \( |y_{k_0}/\sqrt{T}| = O_p(1) \) and \( (1/\sqrt{T}) \sup_{t>k_0} |e_t| = O_p(1) \), the first term is bounded by \( [1/(1 - \beta_2)] |y_{k_0}/\sqrt{k_0}((1/\sqrt{T}) \sup_{t>k_0} |e_t| = o_p(1) \).

Further, because \( E(\sum_{i=k_0+1}^{i-1} \beta_2^{i-1} e_i e_t) = 0 \) and \( \text{Var}(\sum_{i=k_0+1}^{i-1} \beta_2^{i-1} e_i e_t) \leq \sigma^2/(1 - \beta_2^2) < \infty \), by the uniform weak law of large numbers for dependent and heterogeneous processes (Pötscher and Prucha, 1989; Andrews, 1987), the second term converges to zero. This shows (5i).
To show (5j), by (5b), (5i), and the triangle inequality, we have

\[
\sup_{\tau \in \Xi_T} \left| \frac{1}{T} \sum_{[\tau\tau] + 1}^{T} y_{t-1} \varepsilon_t \right| \leq \left| \frac{1}{T} \sum_{[\tau\tau] + 1}^{T} y_{t-1} \varepsilon_t \right| + \sup_{\tau \in \Xi_T} \left| \frac{1}{T} \sum_{[\tau\tau] + 1}^{T} y_{t-1} \varepsilon_t \right| = o_p(1).
\]

(5k) \quad \sup_{\tau \in \Xi_T^I} |\hat{\beta}_1(\tau) - 1| = \sup_{\tau \in \Xi_T^I} \left| \frac{(\beta_2 - 1) \sum_{[\tau\tau] + 1}^{T} y_{t-1}^2 + \sum_{1}^{[\tau\tau]} \varepsilon_t y_{t-1} + \sum_{[\tau\tau] + 1}^{T} \varepsilon_t y_{t-1}}{\sum_{1}^{[\tau\tau]} y_{t-1}^2} \right|

\leq \frac{|\beta_2 - 1| \sum_{[\tau\tau] + 1}^{T} y_{t-1}^2 + \sum_{1}^{[\tau\tau]} \varepsilon_t y_{t-1}}{\sum_{1}^{[\tau\tau]} y_{t-1}^2} + \sup_{\tau \in \Xi_T^I} \left| \frac{\sum_{[\tau\tau] + 1}^{T} \varepsilon_t y_{t-1}}{T} \right|

= O_p(T) + o_p(T) + o_p(T) = O_p\left(\frac{1}{T}\right)

by (5a), (5c), (5d), and (5i).

(5l) \quad \sup_{\tau \in \Xi_T^I} |\hat{\beta}_2(\tau) - \beta_2| = \sup_{\tau \in \Xi_T^I} \left| \frac{\sum_{[\tau\tau] + 1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau\tau] + 1}^{T} y_{t-1}^2} \right| \leq \frac{T}{\sum_{[\tau\tau] + 1}^{T} y_{t-1}^2} \sup_{\tau \in \Xi_T^I} \left| \frac{\sum_{[\tau\tau] + 1}^{T} \varepsilon_t y_{t-1}}{T} \right|

= O_p(1) o_p(1) = o_p(1)

by (5d) and (5j). These prove Lemma 3.

**APPENDIX I: DERIVATION OF EQUATIONS (16)–(18)**

To show that \((1/T)RSS_T(\tau)\) converges uniformly to a random flat line above \(\sigma^2\) for \(\tau \in \Xi_T^I\), it is sufficient to show that

(i) \((1/T)RSS_T(\tau)\) converges pointwise to a random value above \(\sigma^2\) for all \(\tau \in \Xi_T^I\) and

(ii) \(\sup_{\tau_b, \tau_a \in \Xi_T^I} |(1/T)RSS_T(\tau_b) - (1/T)RSS_T(\tau_a)| \xrightarrow{p} 0\) for any \(\tau_b, \tau_a \in \Xi_T^I\).
To show (i), use Lemma 3 and (B.1) in Appendix B. For \( \tau \in \mathbb{Z}^T \), \( \beta_1 = 1 \), and \( |\beta_2| < 1 \), we have

\[
\frac{1}{T} \text{RSS}_T(\tau) = \frac{\sum_{t=1}^{T} e_t^2}{T} - (\hat{\beta}_1(\tau) - 1) \frac{\sum_{t=1}^{[\tau T]} y_{t-1} e_t}{T} - 2(\hat{\beta}_2(\tau) - 1) \frac{\sum_{t=1}^{[\tau T]+1} y_{t-1} e_t}{T}
\]

\[
+ \frac{1}{T} (T(\hat{\beta}_2(\tau) - 1))^2 \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T^2} - \frac{2}{\sqrt{T}} (\hat{\beta}_2(\tau) - \beta_2) \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1} e_t}{\sqrt{T}}
\]

\[
+ (\hat{\beta}_2(\tau) - 1)^2 \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T} + 2(\hat{\beta}_2(\tau) - 1)(1 - \beta_2) \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T}
\]

\[
+ (1 - \beta_2)^2 \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T}
\]

\[
= \frac{\sum_{t=1}^{T} e_t^2}{T} - \sigma_p(1) - \sigma_p(1) + \sigma_p(1) - \sigma_p(1) + \sigma_p(1) + (1 - \beta_2)^2 \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T}
\]

\[
\Rightarrow \sigma^2 + \frac{(1 - \beta_2)(1 - \tau_0 + B^2(\tau_0))}{1 + \beta_2} \sigma^2 > \sigma^2.
\]

To show (ii), use (B.1) in Appendix B, the triangle inequality, and Lemma 3. We have

\[
\sup_{\tau_b, \tau_d \in \mathbb{Z}^T} \left| \frac{1}{T} \text{RSS}_T(\tau_b) - \frac{1}{T} \text{RSS}_T(\tau_d) \right|
\]

\[
= \sup_{\tau_b, \tau_d \in \mathbb{Z}^T} \left| - (\hat{\beta}_1(\tau_b) - 1) \frac{\sum_{t=1}^{[\tau T]} y_{t-1} e_t}{T} - 2(\hat{\beta}_2(\tau_b) - 1) \frac{\sum_{t=1}^{[\tau T]+1} y_{t-1} e_t}{T}
\]

\[
+ (\hat{\beta}_2(\tau_b) - 1)^2 \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T} + 2(\hat{\beta}_2(\tau_b) - \beta_2) \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1} e_t}{T}
\]

\[
+ ((\hat{\beta}_2(\tau_b) - \beta_2)^2 - (\hat{\beta}_2(\tau_d) - \beta_2)^2) \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T} + (\hat{\beta}_1(\tau_d) - 1) \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T}
\]

\[
+ 2(\hat{\beta}_2(\tau_d) - 1) \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T} - (\hat{\beta}_2(\tau_d) - 1)^2 \frac{\sum_{t=[\tau T]+1}^{T} y_{t-1}^2}{T}\right|
\]
This derives equation (16).

For \( \tau = \tau_0 \),

\[
\frac{1}{T} RSS_T(\tau_0) = \frac{T}{T} - \left( \frac{\sum_{t=1}^{[\tau_0 T]} y_{t-1} \varepsilon_t}{T} \right)^2 - 2(\hat{\beta}_2(\tau_0) - \beta_2) \frac{\sum_{t=1}^{[\tau_0 T]} y_{t-1} \varepsilon_t}{T}
\]

\[
+ (\hat{\beta}_2(\tau_0) - \beta_2)^2 \frac{\sum_{t=1}^{[\tau_0 T]} y_{t-1}^2}{T}
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^2 + o_p(1) + o_p(1) + o_p(1) \xrightarrow{p} \sigma^2.
\]

This derives equation (17).

To show that \( \frac{1}{T} RSS_T(\tau) \) converges uniformly to a random linear function above \( \sigma^2 \) for \( \tau \in \mathbb{E}^T_+ \), it is sufficient to show that
(i) $(1/T)\text{RSS}_T(\tau)$ converges pointwise to a random linear function above $\sigma^2$ for all $\tau \in \Xi^T_+$ and

(ii) $\sup_{\tau \in \Xi^T_+} |(1/T)\text{RSS}_T(\tau) - (1/T)\text{RSS}_T(\tau + (1/T))| \xrightarrow{P} 0$ for all $\tau \in \Xi^T_+$.

To show (i), by Lemma 3 and (B.3) in Appendix B, we have

$$\frac{1}{T} \text{RSS}_T(\tau) = \frac{\sum_{t=1}^{T} e_t^2}{T} - 2(\hat{\beta}_1(\tau) - 1) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor} y_{t-1} e_t}{T} + (\hat{\beta}_1(\tau) - 1)^2 \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor} y_{t-1}^2 e_t}{T}$$

$$- 2(\hat{\beta}_1(\tau) - \beta_2) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor + 1} y_{t-1} e_t}{T} + (\hat{\beta}_1(\tau) - \beta_2)^2 \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor + 1} y_{t-1}^2 e_t}{T}$$

$$- (\hat{\beta}_2(\tau) - \beta_2) \frac{\sum_{t=\tau T}^{T} y_{t-1} e_t}{T}$$

$$= \frac{1}{T} \sigma^2 + \frac{(1 - \beta_2)^2 (\tau - \tau_0 + B^2(\tau_0))}{1 - \beta_2^2}$$

To show (ii), use (B.3) in Appendix B, the triangle inequality, and Lemma 3. We have

$$\sup_{\tau \in \Xi^T_+} \left| \frac{1}{T} \text{RSS}_T(\tau) - \frac{1}{T} \text{RSS}_T(\tau + \frac{1}{T}) \right|$$

$$= \sup_{\tau \in \Xi^T_+} \left| 2(\hat{\beta}_1(\tau + \frac{1}{T}) - \hat{\beta}_1(\tau)) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor} y_{t-1} e_t}{T} \right|$$

$$+ \left( (\hat{\beta}_1(\tau) - 1)^2 - (\hat{\beta}_1(\tau + \frac{1}{T}) - 1)^2 \right) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor} y_{t-1}^2 e_t}{T}$$

$$+ 2(\hat{\beta}_1(\tau + \frac{1}{T}) - \beta_2) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor + 1} y_{t-1} e_t}{T} + 2(\hat{\beta}_1(\tau + \frac{1}{T}) - \beta_2) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor + 1} y_{t-1}^2 e_t}{T}$$

$$+ \left( (\hat{\beta}_1(\tau) - \beta_2)^2 - (\hat{\beta}_1(\tau + \frac{1}{T}) - \beta_2)^2 \right) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor + 1} y_{t-1}^2 e_t}{T}$$

$$- (\hat{\beta}_1(\tau + \frac{1}{T}) - \beta_2)^2 \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor} y_{t-1} e_t}{T}$$

$$- (\hat{\beta}_2(\tau + \frac{1}{T}) - \beta_2) \frac{\sum_{t=\tau T}^{\lfloor \tau T \rfloor} y_{t-1}^2 e_t}{T}$$

$$\frac{1}{2} (1 - \beta_2)^2 \frac{\tau - \tau_0 + B^2(\tau_0)}{1 - \beta_2^2}$$
\[
\begin{align*}
&\leq 2 \left| \sum_{t=0}^{[t_0/T]} y_{t-1} \epsilon_t \right| \sup_{\tau \in \mathbb{H}^2} \left| \hat{\beta}_1 (\tau + \frac{1}{T}) - \hat{\beta}_1 (\tau) \right| \\
&+ \frac{1}{T} \left| \sum_{t=1}^{[t_0/T]} y_t^2 \right| \left( \left( T \sup_{\tau \in \mathbb{H}^1} |\hat{\beta}_1 (\tau) - 1| \right)^2 + \left( T \sup_{\tau \in \mathbb{H}^1} \left| \hat{\beta}_1 (\tau + \frac{1}{T}) - 1 \right| \right)^2 \right) \\
&+ 2 \sup_{\tau \in \mathbb{H}^2} \left| \hat{\beta}_1 (\tau + \frac{1}{T}) - \hat{\beta}_1 (\tau) \right| \sup_{\tau \in \mathbb{H}^1} \left| \sum_{t=0}^{[t_0/T]} y_{t-1} \epsilon_t \right| \\
&+ 2 \sup_{\tau \in \mathbb{H}^2} \left| \hat{\beta}_1 (\tau + \frac{1}{T}) - \beta_2 \right| \sup_{\tau \in \mathbb{H}^1} \left| \sum_{t=1}^{[t_0/T]} y_{t} \epsilon_{t+1} \right| \\
&+ \sup_{\tau \in \mathbb{H}^1} \left( \left( \hat{\beta}_1 (\tau) - \beta_2 \right)^2 - \left( \left( \hat{\beta}_1 (\tau + \frac{1}{T}) - \beta_2 \right)^2 \right) \sup_{\tau \in \mathbb{H}^1} \left| \sum_{t=0}^{[t_0/T]} y_{t} \epsilon_{t+1} \right| \\
&+ \sup_{\tau \in \mathbb{H}^1} \left( \left( \hat{\beta}_1 (\tau + \frac{1}{T}) - \beta_2 \right)^2 \sup_{\tau \in \mathbb{H}^1} \left| \sum_{t=1}^{[t_0/T]} \epsilon_{t+1} \right| \\
&+ \sup_{\tau \in \mathbb{H}^1} \left( \hat{\beta}_2 (\tau + \frac{1}{T}) - \hat{\beta}_2 (\tau) \right) \sup_{\tau \in \mathbb{H}^1} \left| \sum_{t=0}^{[t_0/T]} y_{t-1} \epsilon_t \right| \\
&+ \sup_{\tau \in \mathbb{H}^1} \left( \hat{\beta}_2 (\tau + \frac{1}{T}) - \beta_2 \right) \sup_{\tau \in \mathbb{H}^1} \left| \sum_{t=1}^{[t_0/T]} \epsilon_{t+1} \right|
\end{align*}
\]

\( \Rightarrow 0. \)

This derives equation (18).

APPENDIX J:
DERIVATION OF EQUATIONS (19)–(21)

For \( \beta_1 = 1 \), \( |\beta_2| < 1 \) and for \( \tau = \tau_0 - (m/T) \), \( m = 1, 2, 3, \ldots \), and \( (m/T) \to 0 \) as \( T \to \infty \), we have

\[(Ja) \quad \frac{1}{T} y_{k_0-m} \Rightarrow \sigma^2 B^2 (\tau_0) = O_p(1);\]
where

\[
\alpha\left(\tau_0 - \frac{m}{T}\right) = \frac{1}{T} \sum_{k_0-m+1}^{k_0} y_{t-1}^2 = o_p(1), \quad \sum_{k_0+1}^{T} e_t y_{t-1} = o_p(1),
\]

we have

\[
(\text{Jd}) \quad \hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) \Rightarrow \frac{(1 - \beta_2^2) mB^2(\tau_0) + (1 - \tau_0 + B^2(\tau_0))\beta_2}{(1 - \beta_2^2) mB^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)}.
\]

Therefore

\[
\hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) - \beta_2 \Rightarrow \frac{(1 - \beta_2^2) mB^2(\tau_0)(1 - \beta_2)}{(1 - \beta_2^2) mB^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)}.
\]

This derives equation (19).
Using (B.1) in Appendix B, (5a), (5b), and (5d) in Lemma 3, and (Ja)-(Jd), we have

\[
\frac{1}{T} \text{RSS}_T \left( \tau_0 - \frac{m}{T} \right) = \sum_{i=1}^{T} \left( \frac{k_{0-m}}{T} \sum_{j=1}^{k_{0-m}} \frac{y_{i-1} e_i}{y_{i-1}^2} \right)^2 - 2 \left( \frac{\beta_2 \left( \tau_0 - \frac{m}{T} \right) - 1}{T} \right) \sum_{j=1}^{k_{0-m+1}} \frac{y_{i-1}^2}{T} \\
+ \left( \frac{\beta_2 \left( \tau_0 - \frac{m}{T} \right) - 1}{T} \right)^2 \sum_{j=1}^{k_{0-m+1}} \frac{y_{i-1}^2}{T} \\
- 2 \left( \frac{\beta_2 \left( \tau_0 - \frac{m}{T} \right) - \beta_2}{T} \right) \sum_{j=1}^{T} \frac{y_{i-1} e_i}{T} + \left( \frac{\beta_2 \left( \tau_0 - \frac{m}{T} \right) - \beta_2}{T} \right)^2 \sum_{j=1}^{T} \frac{y_{i-1}^2}{T} \\
= \frac{1}{T} + o_p(1) + \frac{1}{T} \sum_{j=1}^{k_{0-m}} \frac{y_{i-1} e_i}{T} + o_p(1) \\
+ \left( \frac{\beta_2 \left( \tau_0 - \frac{m}{T} \right) - \beta_2}{T} \right)^2 \sum_{j=1}^{T} \frac{y_{i-1}^2}{T} \\
\Rightarrow \sigma^2 + \frac{(1 - \tau_0 + B^2(\tau_0))(1 - \beta_2^2)2m\sigma^2B^2(\tau_0)}{(1 - \beta_2^2)mB^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)}.
\]

This derives equation (20).

For \( \tau = \tau_0 + (m/T) \), we have

\[(Je) \quad \sup_{m \in \{1, 2, \ldots, m(T)\} \rightarrow 0} \frac{1}{T} \sum_{j=1}^{k_{0+m}} \frac{y_{i-1} e_i}{T} = o_p(1);\]

\[(Jf) \quad \frac{1}{T} \sum_{j=1}^{k_{0+m}} y_{i-1}^2 \Rightarrow \frac{\sigma^2B^2(\tau_0)(1 - \beta_2^2m)}{1 - \beta_2^2};\]

by (5d) in Lemma 3;

\[(Jg) \quad \sup_{m \in \{1, 2, \ldots, m(T)\} \rightarrow 0} \frac{1}{T} \sum_{j=1}^{T} \frac{y_{i-1} e_i}{T} = o_p(1);\]

\[(Jh) \quad \frac{1}{T} \sum_{j=1}^{T} y_{i-1}^2 \leq \frac{1}{T} \sum_{j=1}^{T} \frac{y_{i-1}^2}{T} = O_p(1)\]
Using (B.3) in Appendix B, by (5a)–(5c) of Lemma 3 and (J)e to (J)j given previously, we can write the residual sum of squares as follows:

\[
\frac{1}{T} \text{RSS}_T\left(\tau_0 + \frac{m}{T}\right) = \frac{\sum_{t=1}^{T} \varepsilon_t^2}{T} - 2\left(\hat{\beta}_1\left(\tau_0 + \frac{m}{T}\right) - 1\right) \sum_{t=1}^{k_0} y_{t-1} \varepsilon_t
\]

\[
+ \left(\hat{\beta}_1\left(\tau_0 + \frac{m}{T}\right) - 1\right)^2 \sum_{t=k_0+1}^{k_0+m} y_{t-1}^2
\]

\[
- 2\left(\hat{\beta}_1\left(\tau_0 + \frac{m}{T}\right) - \beta_2\right) \sum_{t=k_0+1}^{k_0+m} \frac{y_{t-1} \varepsilon_t}{T} + \left(\hat{\beta}_1\left(\tau_0 + \frac{m}{T}\right) - \beta_2\right)^2
\]

\[
\times \frac{\sum_{t=k_0+1}^{T} y_{t-1}^2}{T} - \left(\sum_{t=k_0+1}^{T} y_{t-1} \varepsilon_t\right)^2
\]

\[
= \frac{\sum_{t=1}^{T} \varepsilon_t^2}{T} + o_p(1) + o_p(1) + o_p(1) + \left(\hat{\beta}_1\left(\tau_0 + \frac{m}{T}\right) - \beta_2\right)^2 \sum_{t=k_0+1}^{T} y_{t-1}^2
\]

\[
+ o_p(1)
\]

\[
= \sigma^2 + (1 - \beta_2)\sigma^2 B^2(\tau_0)(1 - \beta_2^2 m) + \frac{\sum_{t=k_0+1}^{T} y_{t-1}^2}{1 + \beta_2}.
\]

This derives equation (21).

**APPENDIX K:**

**PROOF OF THEOREM 4**

It is not difficult to show that \(\hat{\tau}_T\) is \(T\)-consistent by using a similar proof as in Appendices C and G. We want to prove a stronger result that

\[
\lim_{T \to \infty} \Pr(\hat{k} \neq k_0) = 0.
\]
Because \( \hat{k} \leq k_0 + O_p(1) \), for any \( \eta > 0 \), there exists an \( M < \infty \) such that \( \Pr(|\hat{k} - k_0| > M) < \eta \). Therefore,

\[
\Pr(\hat{k} \neq k_0) = \Pr(|\hat{k} - k_0| > M) + \Pr(|\hat{k} - k_0| \leq M \text{ and } \hat{k} \neq k_0)
\]

\[
\leq \eta + \Pr(|\hat{k} - k_0| \leq M \text{ and } \hat{k} \neq k_0)
\]

\[
\leq \eta + \sum_{m=1}^{M} \Pr\left(RSS_T\left(\tau_0 - \frac{m}{T}\right) - RSS_T(\tau_0) < 0\right)
\]

\[
+ \sum_{m=1}^{M} \Pr\left(RSS_T\left(\tau_0 + \frac{m}{T}\right) - RSS_T(\tau_0) < 0\right).
\]

Let \( h_1(m) \) and \( h_2(m) \) be defined as in equations (20) and (21), respectively. The preceding probability is equal to

\[
\eta + \sum_{m=1}^{M} \Pr(h_1(m) + \Delta_{1T} < 0) + \sum_{m=1}^{M} \Pr(h_2(m) + \Delta_{2T} < 0),
\]

where

\[
\Delta_{1T} = RSS_T\left(\tau_0 - \frac{m}{T}\right) - RSS_T(\tau_0) - h_1(m) = o_p(1),
\]

\[
\Delta_{2T} = RSS_T\left(\tau_0 + \frac{m}{T}\right) - RSS_T(\tau_0) - h_2(m) = o_p(1).
\]

Because \( h_1(m) \) and \( h_2(m) \) are increasing with \( m \), this implies every individual probability given earlier is declining with \( m \). Thus, the preceding term is bounded by \( \eta + M \Pr(h_1(1) + \Delta_{1T} < 0) + M \Pr(h_2(1) + \Delta_{2T} < 0) \).

Further, because \( h_1(1) \) and \( h_2(1) \) are positive and of order \( O_p(1) \), there exists a \( T \) large enough such that \( \Pr(h_1(1) + \Delta_{1T} < 0) \leq \eta \) and \( \Pr(h_2(1) + \Delta_{2T} < 0) \leq \eta \) for any \( \eta > 0 \).

Therefore for any \( \eta > 0 \), \( \Pr(\hat{k} \neq k_0) \leq (2M + 1)\eta \) for all large \( T \).

Because \( M \) is finite, we have \( \lim_{T \to \infty} \Pr(\hat{k} = k_0) = 1 \) and \( \lim \hat{\beta}_1(\hat{\tau}_T) = \lim \hat{\beta}_1(\tau_0) = \beta_1 = 1 \), \( \lim \hat{\beta}_2(\hat{\tau}_T) = \lim \hat{\beta}_2(\tau_0) = \beta_2 \). These prove the consistency of \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \).

To find the limiting distributions of \( \hat{\beta}_1(\hat{\tau}_T) \) and \( \hat{\beta}_2(\hat{\tau}_T) \), note that for any given real value \( x \),

\[
\lim_{T \to \infty} \Pr(T(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \leq x)
\]

\[
= \lim_{T \to \infty} \Pr(T(\hat{\beta}_1(\tau_0) - \beta_1) \leq x) \lim_{T \to \infty} \Pr(\hat{k} = k_0)
\]

\[
+ \lim_{T \to \infty} \Pr(T(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \leq x | \hat{k} \neq k_0) \lim_{T \to \infty} \Pr(\hat{k} \neq k_0)
\]

\[
= \lim_{T \to \infty} \Pr(T(\hat{\beta}_1(\tau_0) - \beta_1) \leq x).
\]

Thus, \( \hat{\beta}_1(\hat{\tau}_T) \) and \( \hat{\beta}_1(\tau_0) \) have the same asymptotic distribution. Similarly, \( \hat{\beta}_2(\hat{\tau}_T) \) and \( \hat{\beta}_2(\tau_0) \) have the same asymptotic distribution. By (5a)–(5d) in Lemma 3, we have
where $B(\cdot)$ and $\bar{B}(\cdot)$ are defined in Section 2. These show the limiting distributions of $\hat{\beta}_1(\hat{\tau}_T)$ and $\hat{\beta}_2(\hat{\tau}_T)$ for a fixed magnitude of break.

These derive the limiting distribution of $\hat{\beta}_1$ and $\hat{\beta}_2$ for fixed magnitude of break.

To derive the limiting distribution of $\hat{\tau}_T$ for shrinking break, we fix $\hat{\beta}_1$ at one and let $\beta_{2T} = 1 - [1/\sqrt{T}g(T)]$, where $g(T) > 0$, with $g(T) \to \infty$ and $g(T)/\sqrt{T} \to 0$ as $T \to \infty$. Let $v$ be a finite constant and $B(\cdot), B_1(\cdot), \text{and } B_2(\cdot)$ be defined as in Section 2.

For $\tau = \tau_0 + v[g(T)/T]$ and $v \leq 0$, we have

(i) $\Lambda_T(\tau) = o_p(1)$;

(ii) $(\beta_{2T} - 1)$

\[
\sum_{[\tau T] + 1}^{[\tau_0 T] + 1} y_{\tau -1}^2 \sum_{[\tau T] + 1}^{[\tau_0 T] + 1} y_{\tau -1} \epsilon_{\tau} - \frac{O_p(Tg(T))O_p(T)}{\sqrt{Tg(T)}O_p(T^2)} = O_p\left( \frac{\sqrt{g(T)}}{T} \right)
\]

$= o_p(1)$;

(iii) $\frac{\sum_{[\tau T] + 1}^{[\tau_0 T] + 1} y_{\tau -1}^2}{T} \to 1$;

(iv) $\frac{1}{\sqrt{Tg(T)}} \sum_{0}^{[v g(T)] - 1} y_{k_0 - t - 1} \epsilon_{k_0 - t}$

\[
= \frac{1}{\sqrt{Tg(T)}} \sum_{0}^{[v g(T)] - 1} \left( y_{k_0} - \sum_{i=0}^{t} \epsilon_{k_0 - i} \right) \epsilon_{k_0 - t}
\]

$= \frac{y_{k_0}}{\sqrt{T}} \frac{1}{\sqrt{g(T)}} \sum_{0}^{[v g(T)] - 1} \epsilon_{k_0 - t} - \frac{g(T)}{\sqrt{T}}
\]

$\times \sum_{0}^{[v g(T)] - 1} \left( \frac{1}{\sqrt{g(T)}} \sum_{i=0}^{t} \epsilon_{k_0 - i} \epsilon_{k_0 - t} \right) \frac{1}{g(T)}$

$= - \frac{y_{k_0}}{\sqrt{T}} \frac{1}{\sqrt{g(T)}} \sum_{0}^{[v g(T)] - 1} (-\epsilon_{k_0 - t}) - O_p\left( \frac{g(T)}{\sqrt{T}} \right) \Rightarrow -\sigma^2 B(\tau_0)B_1(-v)$;
Thus, equation (B.2) in Appendix B becomes

\[ RSS_T(\tau) - RSS_T(\tau_0) \]

\[ \begin{align*}
&= -2(\beta_{2T} - 1) \sum_{[\tau T] + 1} y_{i-1} \epsilon_t + (\beta_{2T} - 1)^2 \sum_{[\tau T] + 1} y_{i-1}^2 + o_p(1) \\
&= -\frac{2}{\sqrt{Tg(T)}} \sum_{0}^{[\tau T]} (-y_{k_{0}-i} \epsilon_{k_{0}-i}) + \frac{1}{g(T)} \sum_{0}^{[\nu g(T)]-1} y_{k_{0}-i}^2 + o_p(1) \\
&\Rightarrow -2\sigma^2 B(\tau_0) B_2(\nu) + |\nu|\sigma^2 B^2(\tau_0).
\end{align*} \]

Similarly, for \( \tau = \tau_0 + [\nu g(T)/T] \) and \( \nu > 0 \), by (B.4) in Appendix B and the facts that

\( \Lambda_T(\tau) = o_p(1) \),

\( (\beta_{2T} - 1) \sum_{[\tau T]} y_{i-1} \epsilon_t \sum_{[\tau T] + 1} y_{i-1}^2 = -\frac{O_p(T)O_p(Tg(T))}{\sqrt{Tg(T)O_p(T^2)}} = o_p(1) \),

\( \sum_{1}^{[\tau T]} y_{i-1}^2 \rightarrow 1 \),

equation (B.4) in Appendix B becomes

\[ RSS_T(\tau) - RSS_T(\tau_0) \]

\[ \begin{align*}
&= 2(\beta_{2T} - 1) \sum_{[\tau T] + 1} y_{i-1} \epsilon_t + (\beta_{2T} - 1)^2 \sum_{[\tau T] + 1} y_{i-1}^2 + o_p(1) \\
&= -\frac{2}{\sqrt{Tg(T)}} \sum_{0}^{[\nu g(T)]-1} y_{k_{0}+t} \epsilon_{k_{0}+t} + \frac{1}{Tg(T)} \sum_{0}^{[\nu g(T)]-1} y_{k_{0}+t}^2 + o_p(1) \\
&\Rightarrow -2\sigma^2 B(\tau_0) B_2(\nu) + |\nu|\sigma^2 B^2(\tau_0).
\end{align*} \]
Applying the continuous mapping theorem for argmax functionals, we have

\[(1 - \beta_{2T})^2 T^2 (\hat{\tau} - \tau_0) = \frac{T}{g(T)} (\hat{\tau} - \tau_0) = \hat{v} = \arg \min_{\nu \in \mathbb{R}} \{RSS_T(\tau) - RSS_T(\tau_0)\}\]

\[\frac{d}{d\nu} \arg \min_{\nu \in \mathbb{R}} \left\{ -2\sigma^2 B^2(\tau_0) \left( \frac{B^*(v)}{B(\tau_0)} - \frac{1}{2} |v| \right) \right\} \]

\[= \arg \max_{\nu \in \mathbb{R}} \left( \frac{B^*(v)}{B(\tau_0)} - \frac{1}{2} |v| \right),\]

where \(B^*(v)\) is a two-sided Brownian motion on \(R\) defined to be \(B^*(v) = B_1(-v)\) for \(v \leq 0\) and \(B^*(v) = B_2(v)\) for \(v > 0\).
Structural Change in AR(1) Models
Terence Tai-Leung Chong
Stable URL:
http://links.jstor.org/sici?sici=0266-4666%28200102%2917%3A1%3C87%3ASCIAM%3E2.0.CO%3B2-B

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

**Parameter Inference for a Nearly Nonstationary First-Order Autoregressive Model**
J. Ahtola; G. C. Tiao
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28198408%2971%3A2%3C263%3APIFANN%3E2.0.CO%3B2-H

**Consistency in Nonlinear Econometric Models: A Generic Uniform Law of Large Numbers**
Donald W. K. Andrews
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198711%2955%3A6%3C1465%3ACINEMA%3E2.0.CO%3B2-G

**Laws of Large Numbers for Dependent Non-Identically Distributed Random Variables**
Donald W. K. Andrews
Stable URL:
http://links.jstor.org/sici?sici=0266-4666%28198812%294%3A3%3C458%3ALOLNFD%3E2.0.CO%3B2-9

**Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence**
Anindya Banerjee; Robin L. Lumsdaine; James H. Stock
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C271%3ARASTOT%3E2.0.CO%3B2-C
Asymptotic Inference for Nearly Nonstationary AR(1) Processes
N. H. Chan; C. Z. Wei
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28198709%2915%3A3%3C1050%3AAIFNNA%3E2.0.CO%3B2-5

Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Processes
N. H. Chan; C. Z. Wei
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28198803%2916%3A1%3C367%3ALDOLSE%3E2.0.CO%3B2-Q

Distribution of the Estimators for Autoregressive Time Series With a Unit Root
David A. Dickey; Wayne A. Fuller
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197906%2974%3A366%3C427%3ADOTEFA%3E2.0.CO%3B2-3

Tests for Parameter Instability in Regressions with I(1) Processes
Bruce E. Hansen
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199207%2910%3A3%3C321%3ATFPIIR%3E2.0.CO%3B2-K

Inference About the Change-Point in a Sequence of Random Variables
David V. Hinkley
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28197004%2957%3A1%3C1%3AIATCIA%3E2.0.CO%3B2-9

Cube Root Asymptotics
Jeankyung Kim; David Pollard
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28199003%2918%3A1%3C191%3ACRA%3E2.0.CO%3B2-A
Fixed Accuracy Estimation of an Autoregressive Parameter
T. L. Lai; D. Siegmund
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28198306%2911%3A2%3C478%3AFAEOTA%3E2.0.CO%3B2-3

The Changing Behavior of the Term Structure of Interest Rates
N. Gregory Mankiw; Jeffrey A. Miron
Stable URL:
http://links.jstor.org/sici?sici=0033-5533%28198605%29101%3A2%3C211%3ATCBOTT%3E2.0.CO%3B2-%23

The Adjustment of Expectations to a Change in Regime: A Study of the Founding of the Federal Reserve
N. Gregory Mankiw; Jeffrey A. Miron; David N. Weil
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198706%2977%3A3%3C358%3ATAOETA%3E2.0.CO%3B2-P

Time Series Regression with a Unit Root
P. C. B. Phillips
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198703%2955%3A2%3C277%3ATSRWAU%3E2.0.CO%3B2-R

Regression Theory for Near-Integrated Time Series
P. C. B. Phillips
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198809%2956%3A5%3C1021%3ARTFNTS%3E2.0.CO%3B2-D

Testing and Estimating Change-Points in Time Series
Dominique Picard
Stable URL:
http://links.jstor.org/sici?sici=0001-8678%28198512%2917%3A4%3C841%3ATAECP%3E2.0.CO%3B2-U
A Uniform Law of Large Numbers for Dependent and Heterogeneous Data Processes
Benedikt M. Pötscher; Ingmar R. Prucha
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198905%2957%3A3%3C675%3AAULOLN%3E2.0.CO%3B2-Q

The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case
John S. White
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195812%2929%3A4%3C1188%3ATLDOTS%3E2.0.CO%3B2-P

Approximating the Distribution of the Maximum Likelihood Estimate of the Change-Point in a Sequence of Independent Random Variables
Yi-Ching Yao
Stable URL:
http://links.jstor.org/sici?sici=0090-5364%28198709%2915%3A3%3C1321%3AAATDOTM%3E2.0.CO%3B2-8