Econometric Theory, **17**, 2001, 87–155. Printed in the United States of America.

STRUCTURAL CHANGE IN AR(1) MODELS

TERENCE TAI-LEUNG CHONG
The Chinese University of Hong Kong

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 β_1 and β_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 β_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

This paper is based on chapter 4 of my Ph.D. dissertation submitted to the University of Rochester. I am grateful to my thesis adviser, Bruce Hansen, for his advice and encouragement. My thanks also go to the editor and three anonymous referees for their helpful comments, which led to a substantial revision. All errors are mine. Address correspondence to: Terence Tai-Leung Chong, Department of Economics, the Chinese University of Hong Kong, Shatin, N.T., Hong Kong; email: b792703@mailserv.cuhk.edu.hk.

© 2001 Cambridge University Press 0266-4666/01 $9.50

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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 β_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 β at an unknown time k_0. We consider the following model:

$$y_t = \beta_1 y_{t-1} 1\{t \leq k_0\} + \beta_2 y_{t-1} 1\{t > k_0\} + \varepsilon_t \qquad (t = 1,2,\ldots,T), \tag{1}$$

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

We let $\tau_0 = k_0/T$ 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) $\varepsilon_t \sim i.i.d.(0,\sigma^2)\ \forall t$, $0 < \sigma^2 < \infty$, and $E(\varepsilon_t^4) < \infty$.

(A3) $\tau_0 \in [\underline{\tau}, \bar{\tau}] \subset (0,1)$.

We let $k = [\tau T]$ where $[\cdot]$ is the greatest integer function and define $\underline{k} = [\underline{\tau} T]$ and $\bar{k} = [\bar{\tau} T]$.

Our interests are to estimate the structural parameters β_1 and β_2 and the time of change τ_0. We estimate the following model:

$$\hat{y}_t = \hat{\beta}_1 y_{t-1} 1\{t \leq [\hat{\tau}_T T]\} + \hat{\beta}_2 y_{t-1} 1\{t > [\hat{\tau}_T T]\}. \tag{2}$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

The estimation method employed here is the least squares method proposed in Bai (1994a, 1994b) and Chong (1995). For any given τ, the ordinary least squares (OLS) estimators are given by

$$\hat{\beta}_1(\tau) = \frac{\sum_{1}^{[\tau T]} y_t y_{t-1}}{\sum_{1}^{[\tau T]} y_{t-1}^2}, \tag{3}$$

$$\hat{\beta}_2(\tau) = \frac{\sum_{[\tau T]+1}^{T} y_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2}, \tag{4}$$

and the change-point estimator satisfies

$$\hat{\tau}_T = \underset{\tau \in (0,1)}{\operatorname{Arg\,min}}\, RSS_T(\tau), \tag{5}$$

where

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

We write $\hat{k} = [\hat{\tau}_T T]$, $\hat{\beta}_1 = \hat{\beta}_1(\hat{\tau}_T)$, and $\hat{\beta}_2 = \hat{\beta}_2(\hat{\tau}_T)$. 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 R_+ and $B(\cdot)$ and $\bar{B}(\cdot)$ 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 τ_0 is unknown and has to be estimated by $\hat{\tau}_T$. To show the consistency of $\hat{\tau}_T$, the usual practice is to show that $(1/T)RSS_T(\tau)$ converges uniformly to a nonstochastic function that has a unique minimum at $\tau = \tau_0$. Thus, we will focus on the asymptotic behavior of $(1/T)RSS_T(\tau)$. The

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

following lemma is useful in deriving the limiting behavior of the criterion function $(1/T)RSS_T(\tau)$ and in proving Theorem 1, which follows.

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

$$\text{(3a)} \quad \sup_{0\le\tau_A\le\tau_B\le 1} \frac{1}{T}\left|\sum_{[\tau_A T]}^{[\tau_B T]} y_{t-1}\varepsilon_t\right| = o_p(1);$$

$$\text{(3b)} \quad \frac{1}{T}\sum_{1}^{[\tau_0 T]} y_{t-1}^2 \xrightarrow{p} \frac{\tau_0\sigma^2}{1-\beta_1^2};$$

$$\text{(3c)} \quad \frac{1}{T}\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 \xrightarrow{p} \frac{(1-\tau_0)\sigma^2}{1-\beta_2^2};$$

$$\text{(3d)} \quad \sup_{\underline{\tau}\le\tau\le\tau_0}\left|\frac{1}{T}\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 - \frac{(\tau_0-\tau)\sigma^2}{1-\beta_1^2}\right| = o_p(1);$$

$$\text{(3e)} \quad \sup_{\underline{\tau}\le\tau\le\tau_0} |\hat{\beta}_1(\tau) - \beta_1| = o_p(1);$$

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

$$\text{(3g)} \quad \sup_{\tau_0<\tau\le\bar{\tau}}\left|\frac{1}{T}\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 - \frac{(\tau-\tau_0)\sigma^2}{1-\beta_2^2}\right| = o_p(1);$$

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

$$\text{(3i)} \quad \sup_{\tau_0<\tau\le\bar{\tau}} |\hat{\beta}_2(\tau) - \beta_2| = o_p(1).$$

Proof. See Appendix A.

Note from (3e) that when $\underline{\tau} \le \tau \le \tau_0$, $\hat{\beta}_1(\tau)$ converges uniformly to β_1. This should be obvious because $\hat{\beta}_1(\tau)$ only utilizes the data generated by the process $y_t = \beta_1 y_{t-1} + \varepsilon_t$. However, from (3f), $\hat{\beta}_2(\tau)$ converges uniformly to a weighted average of β_1 and β_2. The weight depends on the true change point, the true preshift and postshift parameters, and the location of τ. When $\tau_0 < \tau \le \bar{\tau}$, (3h) and (3i) display similar results.

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

$$\sup_{\underline{\tau}\le\tau\le\tau_0}\left|\frac{1}{T}RSS_T(\tau) - \sigma^2 - \frac{(\tau_0-\tau)(1-\tau_0)(\beta_2-\beta_1)^2\sigma^2}{(\tau_0-\tau)(1-\beta_2^2) + (1-\tau_0)(1-\beta_1^2)}\right| = o_p(1), \tag{6}$$

$$\sup_{\tau_0<\tau\le\bar{\tau}}\left|\frac{1}{T}RSS_T(\tau) - \sigma^2 - \frac{\tau_0(\tau-\tau_0)(\beta_2-\beta_1)^2\sigma^2}{\tau_0(1-\beta_2^2) + (\tau-\tau_0)(1-\beta_1^2)}\right| = o_p(1). \tag{7}$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

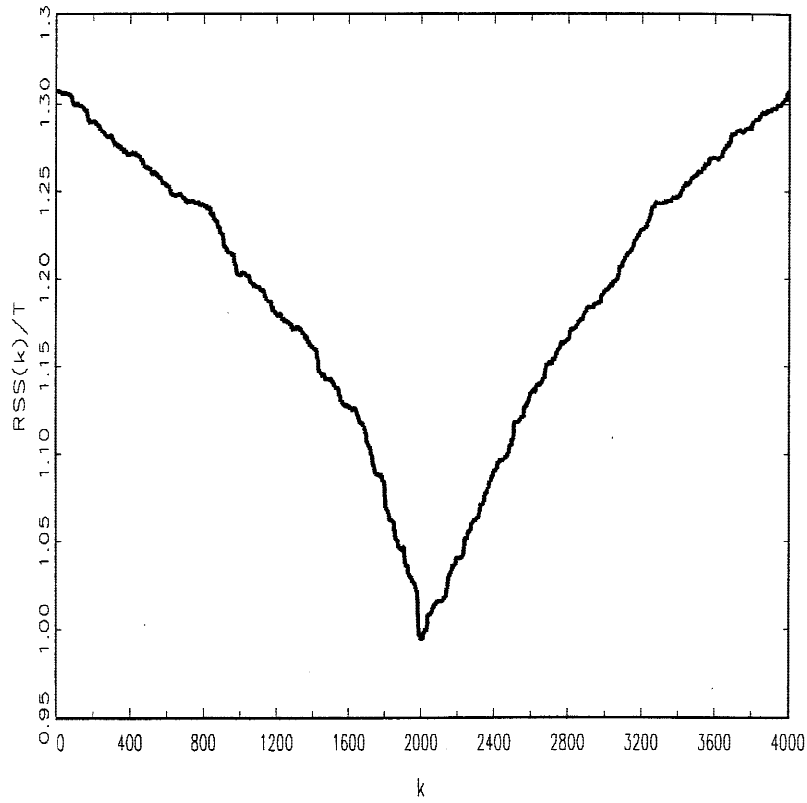


FIGURE 1. Graph of $(1/T)RSS_T(\tau)$ for $\beta_1 = 0.5$, $\beta_2 = -0.5$, $T = 4000$.

One can easily verify that $(1/T)RSS_T(\tau)$ converges uniformly to a piecewise concave function of τ that attains a unique global minimum at $\tau = \tau_0$. A simulation of $(1/T)RSS_T(\tau)$ 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 β_1 and β_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 τ_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*

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$|\hat{\tau}_T - \tau_0| = O_p\left(\frac{1}{T}\right),$$

$$\sqrt{T}(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \xrightarrow{d} N\left(0, \frac{1-\beta_1^2}{\tau_0}\right), \tag{8}$$

$$\sqrt{T}(\hat{\beta}_2(\hat{\tau}_T) - \beta_2) \xrightarrow{d} N\left(0, \frac{1-\beta_2^2}{1-\tau_0}\right).$$

Proof. See Appendix C.

Thus, $\hat{\beta}_1(\hat{\tau}_T)$ and $\hat{\beta}_2(\hat{\tau}_T)$ are asymptotically normally distributed, with variances depending on β_1, β_2, and τ_0.

3.3. The Limiting Distribution of $\hat{\tau}_T$ as β_2 Collapses to β_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 ε_t 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 ε_t, 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 β_1 and let β_{2T} be a sequence of β_2 such that $|\beta_{2T} - \beta_1| \to 0$ and $\sqrt{T}|\beta_{2T} - \beta_1| \to \infty$ as $T \to \infty$.

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

$$\hat{\tau}_T \overset{def}{=} \operatorname*{Arg\,min}_{\tau} RSS_T(\tau) = \operatorname*{Arg\,min}_{\tau}\{RSS_T(\tau) - RSS_T(\tau_0)\}.$$

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

For $v \leq 0$,

$$RSS_T(\tau) - RSS_T(\tau_0) \Rightarrow -2\sigma^2 B_1\left(\frac{|v|}{1-\beta_1^2}\right) + \frac{\sigma^2|v|}{1-\beta_1^2}. \tag{9}$$

For $v > 0$,

$$RSS_T(\tau) - RSS_T(\tau_0) \Rightarrow -2\sigma^2 B_2\left(\frac{v}{1-\beta_1^2}\right) + \frac{\sigma^2 v}{1-\beta_1^2} \tag{10}$$

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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* β_{2T} *be a sequence of* β_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*

$$\frac{T(\beta_{2T} - \beta_1)^2}{1 - \beta_1^2}(\hat{\tau}_T - \tau_0) \xrightarrow{d} \operatorname*{Arg\,max}_{r \in R}\left(B^*(r) - \frac{1}{2}|r|\right), \tag{11}$$

where $B^*(r)$ *is a two-sided Brownian motion on R defined to be* $B^*(r) = B_1(-r)$ *for* $r < 0$ *and* $B^*(r) = B_2(r)$ *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 τ_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 change-point 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 τ_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 $(1/T)RSS_T(\tau)$ 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*

$$\Xi_+^T = (\tau_0, \bar{\tau}] \cap \left\{\tau = \tau_0 + \frac{j}{T} : j = 1, 2, \ldots, [(1 - \tau_0)T] \quad \text{and} \quad \lim_{T\to\infty} \frac{j}{\sqrt{T}} = \infty\right\}.$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Then under Assumptions (A1)–(A3), *we have*

(4a) $\left|\frac{1}{T}\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t\right| = o_p(1);$

(4b) $\frac{1}{T}\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t \Rightarrow \frac{(1-\tau_0)\sigma^2}{2}(B^2(1)-1) = O_p(1);$

(4c) $\frac{1}{T^2}\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 \Rightarrow (1-\tau_0)^2\sigma^2\int_0^1 B^2 = O_p(1);$

(4d) $\sup_{\underline{\tau}\le\tau\le\tau_0}\frac{1}{T}\left|\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t\right| = o_p(1);$

(4e) $\sup_{0\le\tau\le\tau_0}\left|\frac{1}{T}\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 - \frac{(\tau_0-\tau)\sigma^2}{1-\beta_1^2}\right| = o_p(1);$

(4f) $\sup_{\underline{\tau}\le\tau\le\tau_0}|\hat{\beta}_1(\tau)-\beta_1| = o_p(1);$

(4g) $\sup_{\underline{\tau}\le\tau\le\tau_0}|\hat{\beta}_2(\tau)-1| = O_p\left(\frac{1}{T}\right);$

(4h) $\sup_{\underline{\tau}\le\tau\le\tau_0}|\hat{\beta}_2(\tau)-\beta_1| = O_p(1);$

(4i) $\frac{1}{T}\left|\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t\right| = O_p(1) \quad \text{for } \tau\in\Xi_+^T;$

(4j) $\frac{1}{T}\left|\sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t\right| = O_p(1) \quad \text{for } \tau\in\Xi_+^T;$

(4k) $|\hat{\beta}_1(\tau)-1| = O_p\left(\frac{1}{\sqrt{T}}\right) \quad \text{for } \tau\in\Xi_+^T;$

(4l) $|\hat{\beta}_2(\tau)-1| = O_p\left(\frac{1}{T}\right) \quad \text{for } \tau\in\Xi_+^T.$

Proof. See Appendix E.

Note that the set Ξ_+^T is different from the set $(\tau_0,\bar{\tau}]$ as Ξ_+^T excludes sequences of τ that converge too fast to τ_0. For example, Ξ_+^T contains all given constants τ such that $\tau_0 < \tau \le \bar{\tau}$. The sequence $\tau_T = \tau_0 + (1/\log T)$ also belongs to Ξ_+^T. However, the sequence $\tau_T = \tau_0 + (1/T)$ is not in Ξ_+^T. From Appendix F, we have

$$\sup_{\underline{\tau}\le\tau\le\tau_0}\left|\frac{1}{T}RSS_T(\tau) - \sigma^2 - \frac{(1-\beta_1)(\tau_0-\tau)\sigma^2}{1+\beta_1}\right| = o_p(1) \tag{12}$$

and

$$\frac{1}{T}RSS_T(\tau) \xrightarrow{p} \sigma^2 + \frac{(1-\beta_1)\tau_0\sigma^2}{1+\beta_1} \quad \text{for } \tau\in\Xi_+^T. \tag{13}$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Note from (12) and (13) that $(1/T)RSS_T(\tau)$ converges uniformly to a downward sloping linear function for $\underline{\tau} \leq \tau \leq \tau_0$ and it converges to a flat line located above σ^2 for $\tau \in \Xi_+^T$. Further, there is an *asymptotic gap* between $(1/T)RSS_T(\tau_0)$ and $(1/T)RSS_T(\tau)$ for $\tau \in \Xi_+^T$. The transitional process begins at τ_0 and completes at the left boundary of Ξ_+^T. Thus the set Ξ_+^T establishes the rate at which $\operatorname{plim}(1/T)RSS_T(\tau)$ jumps from σ^2 to a flat line located above σ^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 $c > 0$,

$$\hat{\beta}_1(\tau_0 + cT^{a-1}) = \theta_T(a,c)\left(\beta_1 + \frac{\sum_1^{k_0} \varepsilon_t y_{t-1}}{\sum_1^{k_0} y_{t-1}^2}\right) + (1 - \theta_T(a,c))\left(1 + \frac{\sum_{k_0+1}^{k_0+[cT^a]} \varepsilon_t y_{t-1}}{\sum_{k_0+1}^{k_0+[cT^a]} y_{t-1}^2}\right),$$

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

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

and

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

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

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

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.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

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

$$\frac{1}{T} RSS_T\left(\tau_0 + \frac{c}{\sqrt{T}}\right) \Rightarrow \sigma^2 + \frac{\tau_0\sigma^2(1-\beta_1)^2}{\tau_0\left(\int_0^c B_3^2\right)^{-1} + 1 - \beta_1^2} \stackrel{def}{=} \sigma^2 + \varphi(c), \qquad \textbf{(14)}$$

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

Figure 2 simulates the behavior of $(1/T)RSS_T(\tau)$ 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 τ_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 β_1 and τ_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 β_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 β_2 at 1 and let β_{1T} be a sequence of β_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 $v \leq 0$,

$$RSS_T(\tau) - RSS_T(\tau_0) \Rightarrow -2\sigma^2 B_1(|v|)B_a(\tfrac{1}{2}) + |v|\sigma^2 B_a^2(\tfrac{1}{2}).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

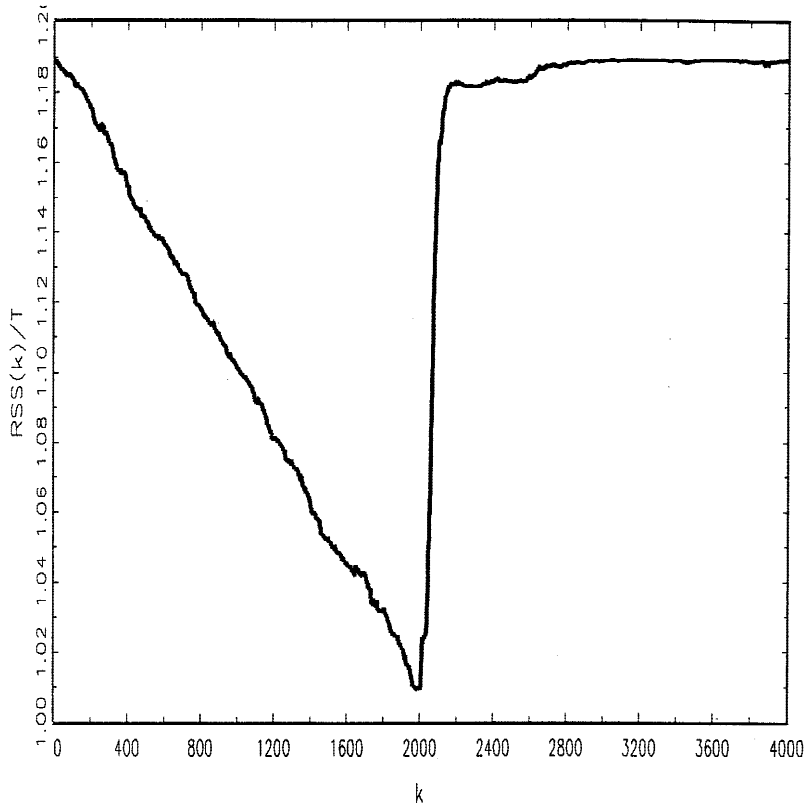


FIGURE 2. Graph of $(1/T)RSS_T(\tau)$ for $\beta_1 = 0.5$, $\beta_2 = 1$, $T = 4000$.

For $v > 0$,

$$RSS_T(\tau) - RSS_T(\tau_0)$$

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

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 β_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*

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$|\hat{\tau}_T - \tau_0| = O_p\left(\frac{1}{T}\right),$$

$$\sqrt{T}(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \xrightarrow{d} N\left(0, \frac{1-\beta_1^2}{\tau_0}\right), \tag{15}$$

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

Suppose we fix β_2 at one and let β_{1T} be a sequence of β_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

$$(1-\beta_{1T})T(\hat{\tau}_T - \tau_0) \xrightarrow{d} \operatorname*{Arg\,max}_{v \in R}\left(\frac{C^*(v)}{B_a\left(\frac{1}{2}\right)} - \frac{1}{2}|v|\right),$$

where $C^*(v)$ is defined to be $C^*(v) = B_1(-v)$ for $v \le 0$ and

$$C^*(v) = -B_2(v) - \int_0^v \frac{B_2(s)}{B_a\left(\frac{1}{2}\right)}\, dB_2(s) - \int_0^v \left(\frac{B_2(s)}{2B_a\left(\frac{1}{2}\right)} + 1\right) B_2(s)\, ds$$

for $v > 0$.

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

Proof. See Appendix G.

An interesting and important feature of Theorem 3 is that the distribution of $\hat{\tau}_T$ is asymmetric about τ_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 τ_0 as shown in Figure 2. As we compress the size of break, this asymmetry perseveres. Thus, $\hat{\tau}_T = \arg\min\ RSS_T(\tau)$ should also be asymmetric about τ_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.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

5.1. The Asymptotic Criterion Function

The following lemma is useful in deriving the limiting behavior of $(1/T)RSS_T(\tau)$ 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*

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

and let Ξ_+^T *be defined as in Lemma* 2. *Under assumptions* (A1)–(A3), *we have*

$$(5a)\quad \frac{1}{T}\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t \Rightarrow \frac{\sigma^2}{2}(B^2(\tau_0) - \tau_0) = O_p(1);$$

$$(5b)\quad \frac{1}{\sqrt{T}}\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t \xrightarrow{d} N\left(0, \frac{\sigma^4}{1-\beta_2^2}\right);$$

$$(5c)\quad \frac{1}{T^2}\sum_{1}^{[\tau_0 T]} y_{t-1}^2 \Rightarrow \sigma^2 \int_0^{\tau_0} B^2 = O_p(1);$$

$$(5d)\quad \frac{1}{T}\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 \Rightarrow \frac{\tau - \tau_0 + B^2(\tau_0)}{1-\beta_2^2}\sigma^2;$$

$$(5e)\quad \sup_{\tau\in\Xi_-^T}\left|\frac{1}{T}\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t\right| = O_p(1);$$

$$(5f)\quad \sup_{\tau\in\Xi_-^T}\left|\frac{1}{T}\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t\right| = O_p(1);$$

$$(5g)\quad \sup_{\tau\in\Xi_-^T}|\hat{\beta}_1(\tau) - 1| = O_p\left(\frac{1}{T}\right);$$

$$(5h)\quad \sup_{\tau\in\Xi_-^T}|\hat{\beta}_2(\tau) - 1| = o_p(1);$$

$$(5i)\quad \sup_{\tau\in\Xi_+^T}\left|\frac{1}{T}\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t\right| = o_p(1);$$

$$(5j)\quad \sup_{\tau\in\Xi_+^T}\left|\frac{1}{T}\sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t\right| = o_p(1);$$

$$(5k)\quad \sup_{\tau\in\Xi_+^T}|\hat{\beta}_1(\tau) - 1| = O_p\left(\frac{1}{T}\right);$$

$$(5l)\quad \sup_{\tau\in\Xi_+^T}|\hat{\beta}_2(\tau) - \beta_2| = o_p(1).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Proof. See Appendix H.

The set Ξ_{-}^{T} serves a similar purpose to that of Ξ_{+}^{T} in Lemma 2. The term Ξ_{-}^{T} excludes sequences of τ that converge to τ_0 at a speed faster than $1/\sqrt{T}$. For example, Ξ_{-}^{T} contains all given constants τ such that $\underline{\tau} \leq \tau < \tau_0$. The sequence $\tau_T = \tau_0 - (1/\log T)$ also belongs to Ξ_{-}^{T}, 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 $\tau \in \Xi_{-}^{T}$,

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

$$\sup_{\tau_a,\tau_b \in \Xi_{-}^{T}} \frac{1}{T} |RSS_T(\tau_b) - RSS_T(\tau_a)| = o_p(1). \tag{16}$$

For $\tau = \tau_0$,

$$\frac{1}{T} RSS_T(\tau_0) \xrightarrow{p} \sigma^2. \tag{17}$$

For $\tau \in \Xi_{+}^{T}$,

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

$$\sup_{\tau \in \Xi_{+}^{T}} \frac{1}{T} \left| RSS_T(\tau) - RSS_T\left(\tau + \frac{1}{T}\right) \right| = o_p(1). \tag{18}$$

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

For $\tau \in \Xi_{+}^{T}$, $(1/T)RSS_T(\tau)$ converges to a random upward-sloping linear function of τ. 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 τ_0, the postshift AR parameter β_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 Ξ_{-}^{T} and Ξ_{+}^{T} in this section to reflect the fact that $\text{plim}(1/T)RSS_T(\tau)$ is discontinuous at both sides of τ_0, whereas we work with the interval $[\underline{\tau}, \tau_0]$ and Ξ_{+}^{T} in Section 4 because $\text{plim}(1/T)RSS_T(\tau)$ is only discontinuous at the right-hand-side neighborhood of τ_0.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Figure 3 simulates the behavior of $(1/T)RSS_T(\tau)$ with $T = 4{,}000$. For $\tau \leq \tau_0$, it is a random flat line and has a sudden plunge near τ_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 $(1/T)RSS_T(\tau_0)$ and $(1/T)RSS_T(\tau_0 \pm c)$. The speed of the transitional process is $1/\sqrt{T}$, and the transition is completed in Ξ_-^T and Ξ_+^T, respectively. To investigate the consistency of the change-point estimator, it is sufficient to show that the asymptotic transition of $(1/T)RSS_T(\tau)$ from $\tau = \tau_0$ to $\tau = \tau_0 \pm c$ is monotonically increasing.

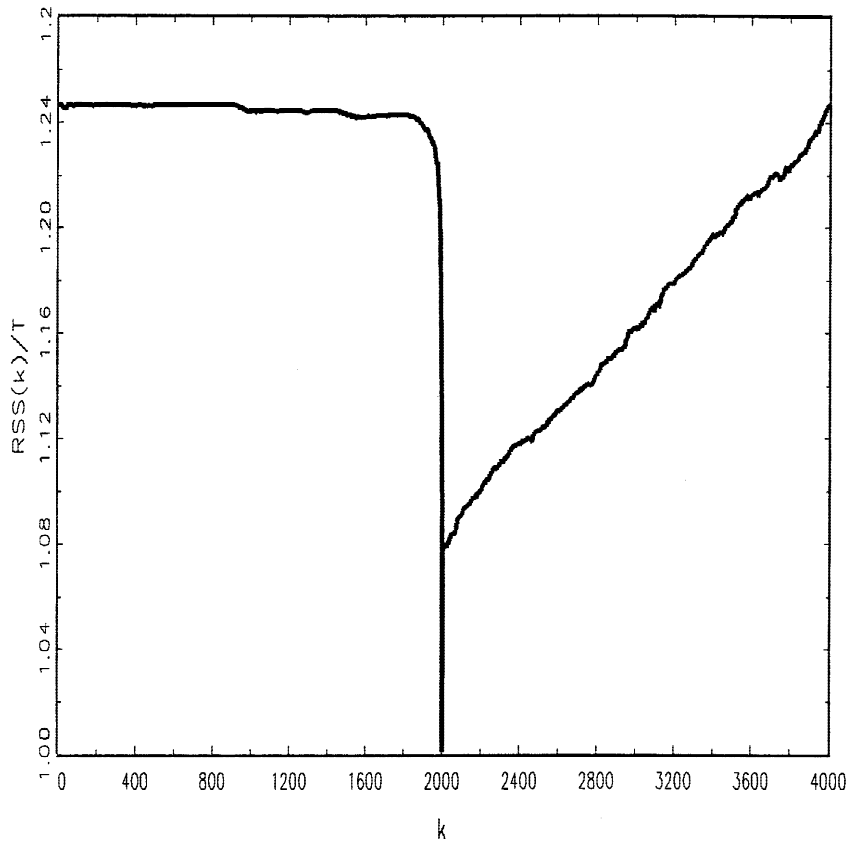


FIGURE 3. Graph of $(1/T)RSS_T(\tau)$ for $\beta_1 = 1$, $\beta_2 = 0.5$, $T = 4000$.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

$$\hat{\beta}_2 \left(\tau_0 - \frac{m}{T} \right) - \beta_2 \Rightarrow \frac{(1-\beta_2^2) m B^2(\tau_0)(1-\beta_2)}{(1-\beta_2^2) m B^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)}, \tag{19}$$

$$\frac{1}{T} RSS_T \left(\tau_0 - \frac{m}{T} \right) \Rightarrow \sigma^2 + \frac{(1-\tau_0+B^2(\tau_0))(1-\beta_2)^2 m B^2(\tau_0)}{(1-\beta_2^2) m B^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)} \sigma^2$$

$$\stackrel{def}{=} \sigma^2 + h_1(m). \tag{20}$$

If $m = 0$, we have $(1/T) RSS_T(\tau_0) \xrightarrow{p} \sigma^2$. As

$$m \to \infty, \frac{1}{T} RSS_T \left(\tau_0 - \frac{m}{T} \right) \Rightarrow \sigma^2 + \frac{(1-\beta_2)(1-\tau_0+B^2(\tau_0))\sigma^2}{1+\beta_2}.$$

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

$$\frac{1}{T} RSS_T \left(\tau_0 + \frac{m}{T} \right) \Rightarrow \sigma^2 + \frac{(1-\beta_2) B^2(\tau_0)(1-\beta_2^{2m})}{1+\beta_2} \sigma^2$$

$$\stackrel{def}{=} \sigma^2 + h_2(m). \tag{21}$$

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 β_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 β_2 in practice. See Theorem 4, which follows.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

Despite the fact that $\hat{\beta}_2$ 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 $T(\hat{\tau}_T - \tau_0) \xrightarrow{p} 0$. Hence the distribution of $\hat{\tau}_T$ degenerates very fast for any fixed magnitude of break. However, if we allow β_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 β_1 at one and let β_{2T} be a sequence of β_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 $\tau = \tau_0 + v/(1-\beta_{2T})^2T^2$, where $v \in R$. From Appendix K, we have the following expressions.

For $v \leq 0$,

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

For $v > 0$,

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

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*

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

$$T(\hat{\beta}_1(\hat{\tau}_T) - 1) \Rightarrow \frac{B^2(\tau_0) - \tau_0}{2\displaystyle\int_0^{\tau_0} B^2}, \tag{22}$$

$$\sqrt{T}(\hat{\beta}_2(\hat{\tau}_T) - \beta_2) \Rightarrow \frac{\sqrt{1-\beta_2^2}\,\bar{B}(1)}{1 - \tau_0 + B^2(\tau_0)}.$$

Suppose we fix β_1 *at one and let* β_{2T} *be a sequence of* β_2 *such that* $\sqrt{T}(1-\beta_{2T}) \to 0$ *and* $T^{3/4}(1-\beta_{2T}) \to \infty$ *as* $T \to \infty$. *Then the limiting distribution of* $\hat{\tau}_T$ *is given by*

$$(1-\beta_{2T})^2T^2(\hat{\tau}_T - \tau_0) \xrightarrow{d} \operatorname*{Arg\,max}_{v\in R}\left(\frac{B^*(v)}{B(\tau_0)} - \frac{1}{2}|v|\right),$$

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* $v > 0$. *The terms* $B_1(\cdot)$, $B_2(\cdot)$, $B(\cdot)$, *and* $\bar{B}(\cdot)$ *are defined in Section* 2.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

$$\sqrt{T}(\hat{\beta}_2 - \beta_2) = \frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}/\sqrt{T}}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2/T} + o_p(1),$$

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 β_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); $\{\varepsilon_t\}_{t=1}^{T} \sim nid(0,1)$ and $y_0 \sim nid(0,1)$ independent of $\{\varepsilon_t\}_{t=1}^{T}$. The true change point is set at $\tau_0 = 0.5$.

Experiment 1a. This experiment verifies equation (19), which predicts that when $\beta_1 = 1$ and $|\beta_2| < 1$, the postshift structural estimator $\hat{\beta}_2(\tau)$ does not converge to β_2 whenever $\tau < \tau_0$. We let $\beta_1 = 1$; $\beta_2 = 0.5, 0, -0.5$.

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

Note that when $\beta_1 = 1$ and $|\beta_2| < 1$, the estimate of β_2 is inaccurate for $\tau < \tau_0$. The strange behavior of $\bar{\hat{\beta}}_2(\tau)$ is highlighted by italic figures in Table 1. Note that $\bar{\hat{\beta}}_2(\tau)$ is not close to β_2 even for $\tau = \tau_0 - (1/T) = 0.495$.

Experiment 1b. With $|\beta_1| < 1$, $\beta_2 = 1$, and τ close to τ_0, we will show that both $\hat{\beta}_1(\tau)$ and $\hat{\beta}_2(\tau)$ are very close to the true parameters β_1 and β_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 $\bar{\hat{\beta}}_1(\tau)$ is very close to β_1. Note that in experiments 1a and 1b, the estimates of the unit root are close to one in all cases.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

TABLE 1. Results for experiment 1a

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

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

$\tau \setminus \bar{\bar{\beta}}$	$\bar{\bar{\beta}}_1(\tau)$ $\beta_1 = 0.5$	$\bar{\bar{\beta}}_2(\tau)$ $\beta_2 = 1$	$\bar{\bar{\beta}}_1(\tau)$ $\beta_1 = 0$	$\bar{\bar{\beta}}_2(\tau)$ $\beta_2 = 1$	$\bar{\bar{\beta}}_1(\tau)$ $\beta_1 = -0.5$	$\bar{\bar{\beta}}_2(\tau)$ $\beta_2 = 1$
0.480	0.4899	0.9817	−0.0001	0.9810	−0.4902	0.9788
0.485	0.4900	0.9820	−0.0000	0.9815	−0.4903	0.9799
0.490	0.4900	0.9823	−0.0000	0.9820	−0.4904	0.9809
0.495	0.4901	0.9827	−0.0000	0.9825	−0.4905	0.9819
0.500	0.4902	0.9830	−0.0000	0.9830	−0.4905	0.9830
0.505	0.4956	0.9830	0.0109	0.9830	−0.4756	0.9830
0.510	0.5044	0.9830	0.0301	0.9830	−0.4512	0.9829
0.515	0.5158	0.9830	0.0559	0.9830	−0.4193	0.9829
0.520	0.5290	0.9829	0.0862	0.9830	−0.3819	0.9829

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

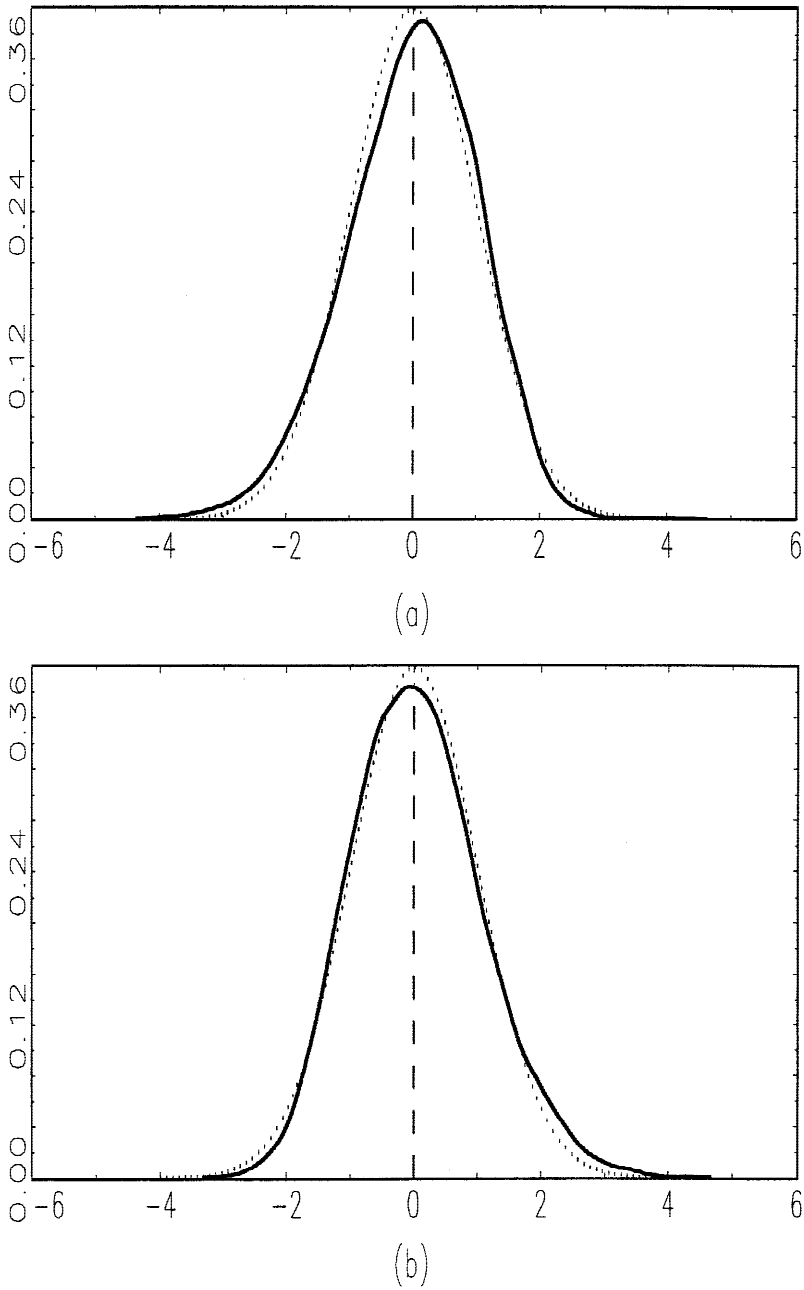


FIGURE 4. (a) Distribution of $\sqrt{\tau_0 T/(1-\beta_1^2)}(\hat{\beta}_1 - \beta_1)$; —— ($T = 200$), ····· ($T = \infty$). (b) Distribution of $\sqrt{(1-\tau_0)T/(1-\beta_2^2)}(\hat{\beta}_2 - \beta_2)$; —— ($T = 200$), ····· ($T = \infty$).

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

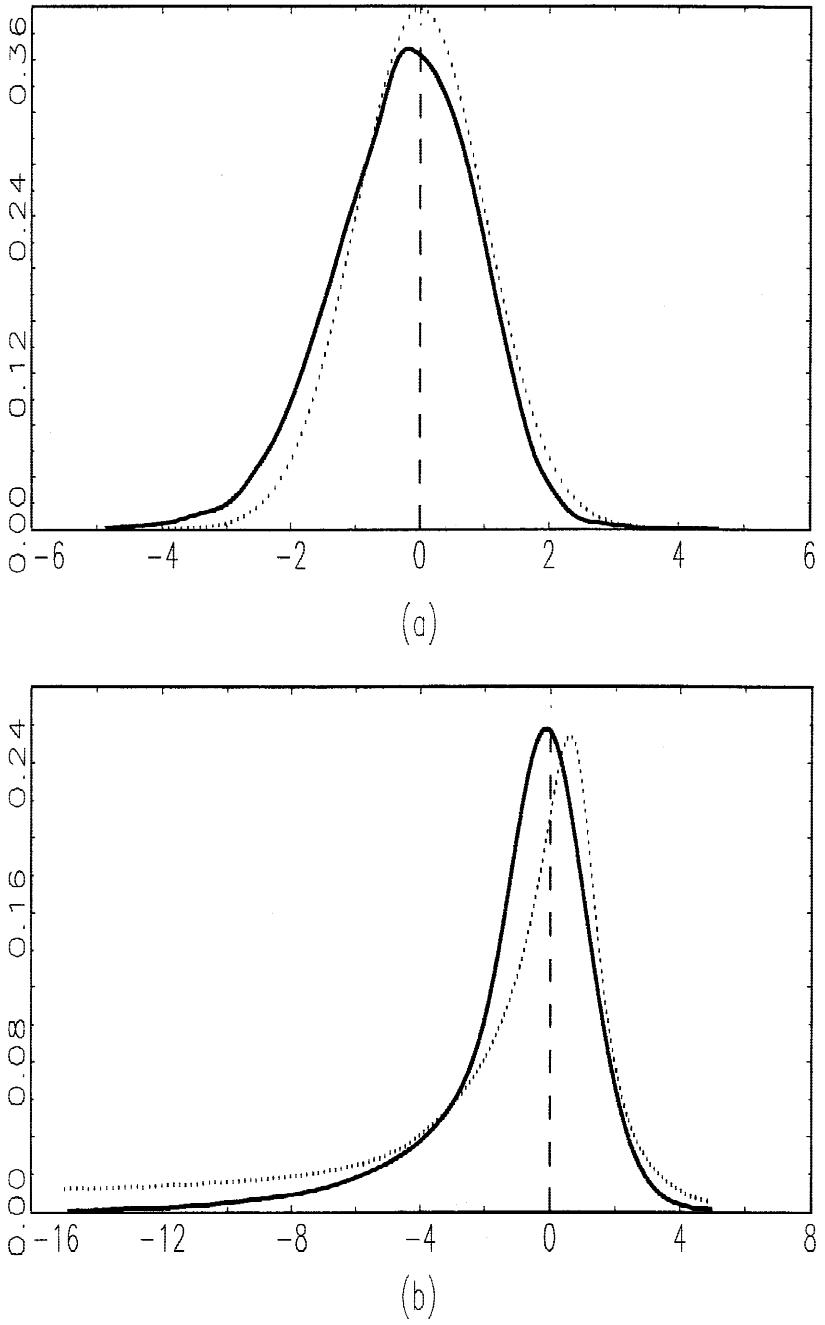


FIGURE 5. (a) Distribution of $\sqrt{\tau_0 T/(1-\beta_1^2)}(\hat{\beta}_1 - \beta_1)$; —— ($T = 200$), ····· ($T = \infty$). (b) Distribution of $(1-\tau_0)T/(\hat{\beta}_2 - \beta_2)$; —— ($T = 200$), ····· ($T = \infty$).

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

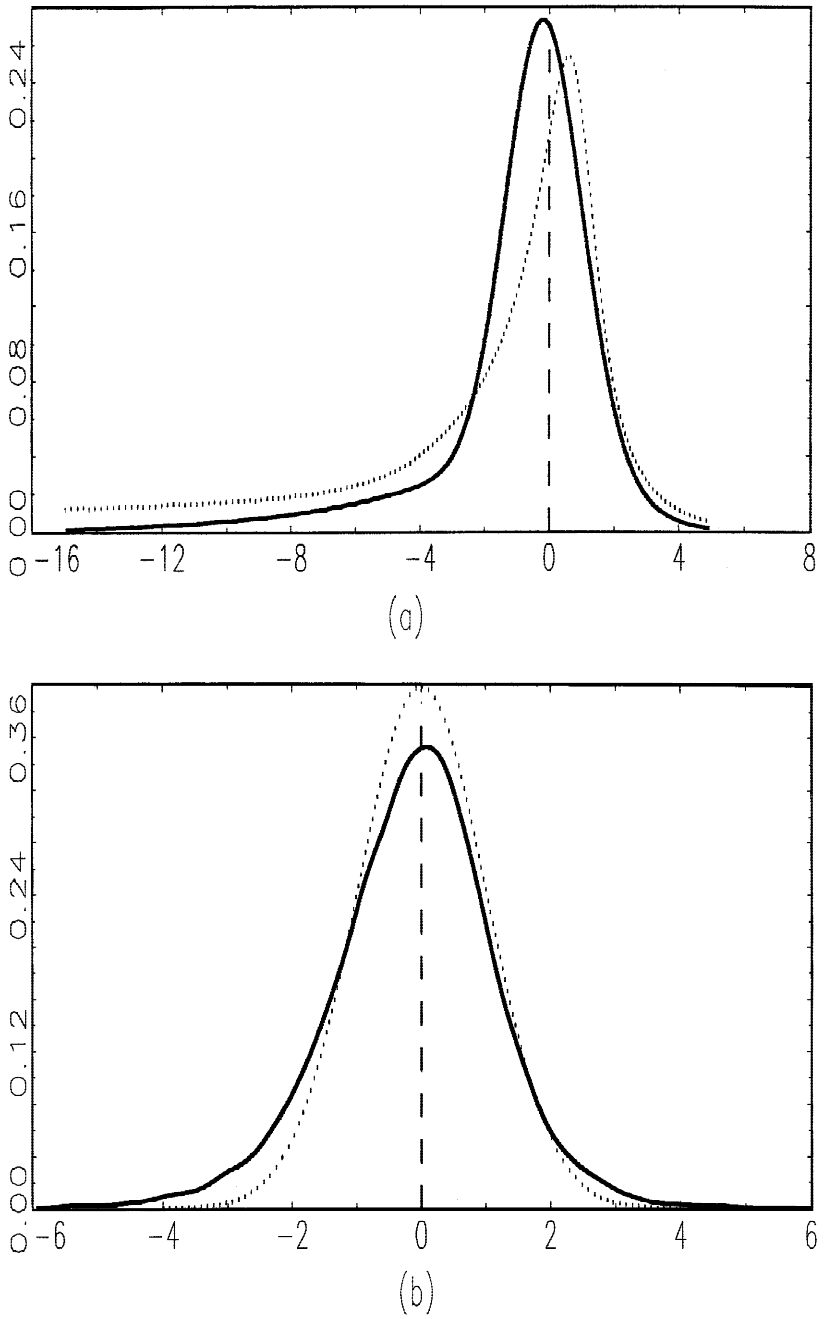


FIGURE 6. (a) Distribution of $\tau_0 T(\hat{\beta}_1 - \beta_1)$; —— ($T = 200$), ····· ($T = \infty$). (b) Distribution of $\sqrt{T/(1-\beta_2^2)}(\hat{\beta}_2 - \beta_2)$; —— ($T = 200$), ····· ($T = \infty$).

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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 $\hat{\tau}_T$ 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 nonstationary 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 $\hat{\tau}_T$ for shrinking break.

Case 1: $\beta_1 = 0.5, \beta_{2T} = 0.5 - T^{-1/8}$;
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 $T^{3/4}(\hat{\tau}_T - \tau_0)$.

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 $\hat{\tau}_T$ is symmetrically distributed, as predicted by Theorems 1 and 4.

For Case 2, Theorem 3 predicts that $\hat{\tau}_T$ should be asymmetrically distributed. Figure 7b agrees with our predictions.

Theoretically, as $T \to \infty$, the domain for $T^{3/4}(\hat{\tau}_T - \tau_0)$ should be unbounded. However, because we fix T at 200 in this experiment, and $\hat{\tau}_T$ is searched within the (0,1) interval, hence the simulated distribution of $T^{3/4}(\hat{\tau}_T - \tau_0)$ 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 $\hat{\tau}_T$ 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 $\mathrm{Var}(\hat{\beta}_2(\tau) - \hat{\beta}_1(\tau))$ when τ 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 τ, 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 τ is near the boundary. For example, when τ 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 $\hat{\tau}_T$ 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 $\hat{\tau}_T$ 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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

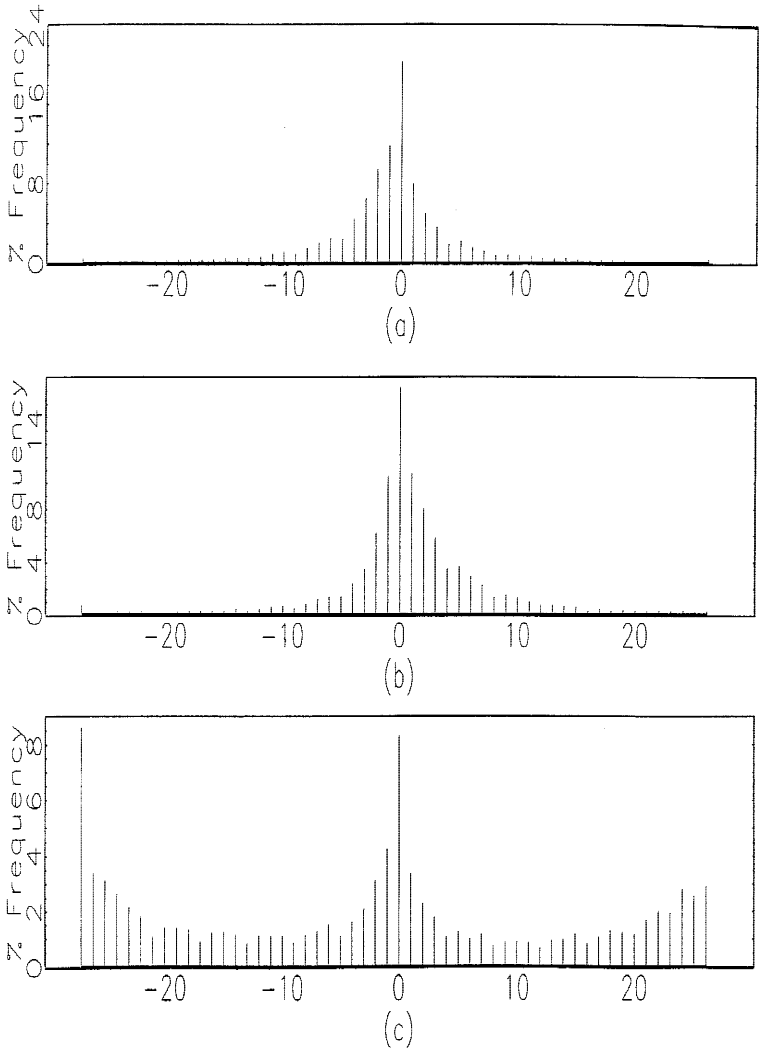


FIGURE 7. Distribution of $T^{3/4}(\hat{\tau}_T - 0.5)$: (a) $\beta_1 = 0.5$, $\beta_2 = 0.5 - T^{-1/8}$; (b) $\beta_1 = 1 - T^{-1/4}$, $\beta_2 = 1$; (c) $\beta_1 = 1$, $\beta_2 = 1 - T^{-5/8}$.

(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-

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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, $y_t = y_{t-1}1\{t \leq k_0\} + (y_{k_0}(1 - \beta_2) + \beta_2 y_{t-1})1\{t > k_0\} + \varepsilon_t$. This process avoids a spurious sharp jump to zero at the break point.

It should also be mentioned that the disturbance terms $\{\varepsilon_t\}_{t=1}^{T}$ 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 ε_t. This assumption is also very helpful in calculating the long run variances of processes such as $\sum \varepsilon_t y_{t-1}$. However, extension to the cases of heterogeneous and/or dependent ε_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

Ahtola J. & G.C. Tiao (1984) Parameter inference for a nearly nonstationary first order autoregressive model. *Biometrika* 71, 263–272.

Andrews, D.W.K. (1987) Consistency in nonlinear econometric models: A generic uniform law of large numbers. *Econometrica* 55, 1465–1471.

Andrews, D.W.K. (1988) Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory* 4, 458–467.

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

Bai, J. (1994b) Estimation of Structural Change Based on Wald Type Statistics. Working paper 94-6, M.I.T., Department of Economics.

Banerjee, A., R.L. Lumsdaine, & J.H. Stock (1992) Recursive and sequential tests of the unit root and trend break hypothesis: Theory and international evidence. *Journal of Business and Economic Statistics* 10, 271–287.

Chan, N.H. & C.Z. Wei (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.

Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least squares estimate of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.

Chong, T.T.L. (1995) Partial parameter consistency in a misspecified structural change model. *Economics Letters* 49, 351–357.

Dickey, D.A. & W.A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Hansen, B.E. (1992) Tests for parameter instability in regressions with I(1) processes. *Journal of Business and Economic Statistics* 10, 321–335.
Hinkley, D.V. (1970) Inference about the change-point in a sequence of random variables. *Biometrika* 57, 1–17.
Kim, J. & D. Pollard (1990) Cube root asymptotics. *Annals of Statistics* 18, 191–219.
Lai, T.L. & D. Siegmund (1983) Fixed accuracy estimation of an autoregressive parameter. *Annals of Statistics* 11, 478–485.
Mankiw, N.G. & J.A. Miron (1986) The changing behavior of the term structure of interest rates. *Quarterly Journal of Economics* 101, 221–228.
Mankiw, N.G., J.A. Miron & D.N. Weil (1987) The adjustment of expectations to a change in regime: A study of the founding of the Federal Reserve. *American Economic Review* 77, 358–374.
Perron, P. (1996) The adequacy of asymptotic approximations in the near-integrated autoregressive model with dependent errors. *Journal of Econometrics* 70, 317–350.
Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–301.
Phillips, P.C.B. (1988) Regression theory for near-integrated time series. *Econometrica* 56, 1021–1043.
Picard, D. (1985) Testing and estimating change-points in time series. *Advances in Applied Probability* 17, 841–867.
Pötscher, B.M. & I.R. Prucha (1989) A uniform law of large numbers for dependent and heterogeneous data process. *Econometrica* 57, 675–683.
Salazar, D. (1982) Structural changes in time series models. *Journal of Econometrics* 19, 147–163.
White, H. (1984) *Asymptotic Theory for Econometricians*. New York: Academic Press.
White, J.S. (1958) The limiting distribution of the serial correlation coefficient in the explosive case. *Annals of Mathematical Statistics* 29, 1188–1197.
Yao, Y.C. (1987) Approximating the distribution of the maximum likelihood estimate of the change-point in a sequence of independent random variables. *Annals of Statistics* 15, 1321–1328.

APPENDIX A: PROOF OF LEMMA 1

If $|\beta_1| < 1$, $|\beta_2| < 1$, ε_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).

$$\text{(3e)} \quad \sup_{\underline{\tau} \le \tau \le \tau_0} |\hat{\beta}_1(\tau) - \beta_1| = \sup_{\underline{\tau} \le \tau \le \tau_0} \left| \frac{\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t}{\sum_{1}^{[\tau T]} y_{t-1}^2} \right|$$

$$\le \frac{T}{\sum_{1}^{[\underline{\tau} T]} y_{t-1}^2} \sup_{\underline{\tau} \le \tau \le \tau_0} \left| \frac{\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t}{T} \right| = O_p(1)o_p(1) = o_p(1).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

To show (3f), utilizing the fact that

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

adding and subtracting

$$(\beta_2 - \beta_1) \frac{T \dfrac{(1-\tau_0)\sigma^2}{1-\beta_2^2}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2},$$

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

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

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

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

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

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

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

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

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

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

The proof for (3h) is analogous to (3f), using the fact that

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

adding and subtracting

$$(\beta_2 - \beta_1) \frac{T \dfrac{(\tau - \tau_0)\sigma^2}{1-\beta_2^2}}{\sum_{1}^{[\tau T]} y_{t-1}^2},$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

by the triangle inequality, and last by (3a), (3b), and (3g) we have

$$\sup_{\tau_0<\tau\le\bar\tau}\left|\hat\beta_1(\tau)-\frac{\tau_0(1-\beta_2^2)\beta_1+(\tau-\tau_0)(1-\beta_1^2)\beta_2}{\tau_0(1-\beta_2^2)+(\tau-\tau_0)(1-\beta_1^2)}\right|$$

$$\le \sup_{\tau_0<\tau\le\bar\tau}\left|\frac{\sum_1^{[\tau T]}\varepsilon_t y_{t-1}}{\sum_1^{[\tau T]}y_{t-1}^2}\right|+|\beta_2-\beta_1|\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{\sum_{[\tau_0 T]+1}^{[\tau T]}y_{t-1}^2}{\sum_1^{[\tau T]}y_{t-1}^2}-\frac{T\dfrac{(\tau-\tau_0)\sigma^2}{1-\beta_2^2}}{\sum_1^{[\tau T]}y_{t-1}^2}\right|$$

$$+|\beta_2-\beta_1|\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{T\dfrac{(\tau-\tau_0)\sigma^2}{1-\beta_2^2}}{\sum_1^{[\tau T]}y_{t-1}^2}-\frac{(\tau-\tau_0)(1-\beta_1^2)}{\tau_0(1-\beta_2^2)+(\tau-\tau_0)(1-\beta_1^2)}\right|$$

$$\le \frac{T}{\sum_1^{[\tau_0 T]}y_{t-1}^2}\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{\sum_1^{[\tau T]}\varepsilon_t y_{t-1}}{T}\right|+\frac{|\beta_2-\beta_1|T}{\sum_1^{[\tau_0 T]}y_{t-1}^2}\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{\sum_{[\tau_0 T]+1}^{[\tau T]}y_{t-1}^2}{T}-\frac{(\tau-\tau_0)\sigma^2}{1-\beta_2^2}\right|$$

$$+\sup_{\tau_0<\tau\le\bar\tau}\frac{|\beta_2-\beta_1|(\tau-\tau_0)(1-\beta_1^2)T}{[\tau_0(1-\beta_2^2)+(\tau-\tau_0)(1-\beta_1^2)]\sum_1^{[\tau T]}y_{t-1}^2}$$

$$\times\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{[\tau_0(1-\beta_2^2)+(\tau-\tau_0)(1-\beta_1^2)]\sigma^2}{(1-\beta_1^2)(1-\beta_2^2)}-\frac{\sum_1^{[\tau T]}y_{t-1}^2}{T}\right|$$

$$\le \frac{T}{\sum_1^{[\tau_0 T]}y_{t-1}^2}\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{\sum_1^{[\tau T]}\varepsilon_t y_{t-1}}{T}\right|+\frac{|\beta_2-\beta_1|T}{\sum_1^{[\tau_0 T]}y_{t-1}^2}\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{\sum_{[\tau_0 T]+1}^{[\tau T]}y_{t-1}^2}{T}-\frac{(\tau-\tau_0)\sigma^2}{1-\beta_2^2}\right|$$

$$+\frac{|\beta_2-\beta_1|T}{\sum_1^{[\tau_0 T]}y_{t-1}^2}\left(\left|\frac{\tau_0\sigma^2}{1-\beta_1^2}-\frac{\sum_1^{[\tau_0 T]}y_{t-1}^2}{T}\right|+\sup_{\tau_0<\tau\le\bar\tau}\left|\frac{(\tau-\tau_0)\sigma^2}{1-\beta_2^2}-\frac{\sum_{[\tau_0 T]+1}^{[\tau T]}y_{t-1}^2}{T}\right|\right)$$

$$=o_p(1).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$\text{(3i)}\quad \sup_{\tau_0<\tau\le\bar{\tau}} |\hat{\beta}_2(\tau)-\beta_2| = \sup_{\tau_0<\tau\le\bar{\tau}} \left| \frac{\sum_{[\tau T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} \right|$$

$$\le \frac{T}{\sum_{[\bar{\tau}T]+1}^{T} y_{t-1}^2} \sup_{\tau_0<\tau\le\bar{\tau}} \left| \frac{\sum_{[\tau T]+1}^{T} \varepsilon_t y_{t-1}}{T} \right| = O_p(1)o_p(1) = o_p(1).$$

These prove Lemma 1. ■

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

For $\underline{\tau} \le \tau \le \tau_0$, we have

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

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

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

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

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

By (B.1) and the fact that

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

we have

$$
\begin{aligned}
&RSS_T(\tau) - RSS_T(\tau_0) \\
&= \frac{\left(\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t\right)^2}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2} - \frac{\left(\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t\right)^2}{\sum_{1}^{[\tau T]} y_{t-1}^2} \\
&\quad - 2\left((\beta_2-\beta_1)\frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} + \frac{\sum_{[\tau T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2}\right)\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t \\
&\quad + \left((\beta_2-\beta_1)\frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} + \frac{\sum_{[\tau T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2}\right)^2 \sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 \\
&\quad - 2\left((\beta_1-\beta_2)\frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} + \frac{\sum_{[\tau T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} - \frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}\right)\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t \\
&\quad + \left[\left((\beta_1-\beta_2)\frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} + \frac{\sum_{[\tau T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2}\right)^2 - \left(\frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}\right)^2\right]\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 \\
&= 2(\beta_2-\beta_1)\left(\frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 \sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} - \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 \sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t}{\sum_{[\tau T]+1}^{T} y_{t-1}^2}\right) \\
&\quad + (\beta_2-\beta_1)^2 \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 \sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} + \Lambda_T(\tau), \qquad \textbf{(B.2)}
\end{aligned}
$$

where

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Let

$$\lambda_1 = \left(\frac{(1-\tau_0)(1-\beta_1^2)(\beta_2-\beta_1)}{(\tau_0-\tau)(1-\beta_2^2)+(1-\tau_0)(1-\beta_1^2)} \right)^2 \quad \text{and}$$

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

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

$$\sup_{\underline{\tau}\le\tau\le\tau_0} \left| \frac{1}{T} RSS_T(\tau) - \sigma^2 - \frac{(\beta_2-\beta_1)^2\sigma^2(\tau_0-\tau)(1-\tau_0)}{(\tau_0-\tau)(1-\beta_2^2)+(1-\tau_0)(1-\beta_1^2)} \right|$$

$$\le \left| \frac{1}{T}\sum_1^T \varepsilon_t^2 - \sigma^2 \right| + \sup_{\underline{\tau}\le\tau\le\tau_0} \frac{\left(\sum_1^{[\tau T]} y_{t-1}\varepsilon_t \right)^2}{T\sum_1^{[\tau T]} y_{t-1}^2}$$

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

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

$$+ \sup_{\underline{\tau}\le\tau\le\tau_0} \left| 2(\hat{\beta}_2(\tau)-\beta_2)\frac{1}{T}\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t \right|$$

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

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 $\sup|a^2-b^2| \le \sup|a-b|\sup|a+b|$, and Lemma 1, the preceding expression is bounded by

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$
\begin{aligned}
&\leq \left| \frac{1}{T}\sum_{1}^{T}\varepsilon_t^2 - \sigma^2 \right| + \sup_{\underline{\tau}\leq\tau\leq\tau_0} |\hat{\beta}_1(\tau)-\beta_1| \sup_{\underline{\tau}\leq\tau\leq\tau_0} \left| \frac{1}{T}\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t \right| \\
&\quad + 2 \sup_{\underline{\tau}\leq\tau\leq\tau_0} |\hat{\beta}_2(\tau)-\beta_1| \sup_{\underline{\tau}\leq\tau\leq\tau_0} \left| \frac{1}{T}\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t \right| \\
&\quad + \sup_{\underline{\tau}\leq\tau\leq\tau_0} (\hat{\beta}_2(\tau)-\beta_1)^2 \sup_{\underline{\tau}\leq\tau\leq\tau_0} \left| \frac{1}{T}\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 - \frac{(\tau_0-\tau)\sigma^2}{1-\beta_1^2} \right| \\
&\quad + \sup_{\underline{\tau}\leq\tau\leq\tau_0} \frac{(\tau_0-\tau)\sigma^2}{1-\beta_1^2} \sup_{\underline{\tau}\leq\tau\leq\tau_0} |(\hat{\beta}_2(\tau)-\beta_1)^2 - \lambda_1| \\
&\quad + 2 \sup_{\underline{\tau}\leq\tau\leq\tau_0} |\hat{\beta}_2(\tau)-\beta_2| \left| \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t}{T} \right| \\
&\quad + \sup_{\underline{\tau}\leq\tau\leq\tau_0} (\hat{\beta}_2(\tau)-\beta_2)^2 \left| \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{T} - \frac{(1-\tau_0)\sigma^2}{1-\beta_2^2} \right| \\
&\quad + \frac{(1-\tau_0)\sigma^2}{1-\beta_2^2} \sup_{\underline{\tau}\leq\tau\leq\tau_0} |(\hat{\beta}_2(\tau)-\beta_2)^2 - \lambda_2| \\
&= o_p(1).
\end{aligned}
$$

This derives equation (6). ■

For $\tau_0 < \tau \leq \bar{\tau}$, we can write the residual sum of squares as follows:

$$
\begin{aligned}
RSS_T(\tau) &= \sum_{1}^{[\tau_0 T]} (\varepsilon_t - (\hat{\beta}_1(\tau)-\beta_1)y_{t-1})^2 + \sum_{[\tau_0 T]+1}^{[\tau T]} (\varepsilon_t - (\hat{\beta}_1(\tau)-\beta_2)y_{t-1})^2 \\
&\quad + \sum_{[\tau T]+1}^{T} (\varepsilon_t - (\hat{\beta}_2(\tau)-\beta_2)y_{t-1})^2 \\
&= \sum_{1}^{T}\varepsilon_t^2 - 2(\hat{\beta}_1(\tau)-\beta_1)\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t + (\hat{\beta}_1(\tau)-\beta_1)^2 \sum_{1}^{[\tau_0 T]} y_{t-1}^2 \\
&\quad - 2(\hat{\beta}_1(\tau)-\beta_2)\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t + (\hat{\beta}_1(\tau)-\beta_2)^2 \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 \\
&\quad - \frac{\left(\sum_{[\tau T]+1}^{T} \varepsilon_t y_{t-1}\right)^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2}.
\end{aligned}
\tag{B.3}
$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

By (B.3) and the fact that

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

we have

$$RSS_T(\tau) - RSS_T(\tau_0) = 2(\beta_2-\beta_1)\left(\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t \sum_{1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{1}^{[\tau T]} y_{t-1}^2} - \frac{\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2}{\sum_{1}^{[\tau T]} y_{t-1}^2}\right)$$
$$+ (\beta_2-\beta_1)^2 \frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 \sum_{1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{1}^{[\tau T]} y_{t-1}^2} + \Lambda_T(\tau). \qquad \textbf{(B.4)}$$

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

$$\sup_{\tau_0<\tau\le\bar{\tau}}\left|\frac{1}{T}RSS_T(\tau) - \sigma^2 - \frac{\tau_0(\beta_2-\beta_1)^2\sigma^2(\tau-\tau_0)}{\tau_0(1-\beta_2^2)+(\tau-\tau_0)(1-\beta_1^2)}\right|$$
$$\le \left|\frac{1}{T}\sum_{1}^{T}\varepsilon_t^2 - \sigma^2\right| + \sup_{\tau_0<\tau\le\bar{\tau}}\left|(\hat{\beta}_1(\tau)-\beta_1)\frac{2}{T}\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t\right|$$
$$+ \sup_{\tau_0<\tau\le\bar{\tau}}\left|(\hat{\beta}_1(\tau)-\beta_1)^2\frac{1}{T}\sum_{1}^{[\tau_0 T]} y_{t-1}^2 - \lambda_3\frac{\tau_0\sigma^2}{1-\beta_1^2}\right|$$
$$+ \sup_{\tau_0<\tau\le\bar{\tau}}\left|(\hat{\beta}_1(\tau)-\beta_2)\frac{2}{T}\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t\right|$$
$$+ \sup_{\tau_0<\tau\le\bar{\tau}}\left|(\hat{\beta}_1(\tau)-\beta_2)^2\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2}{T} - \lambda_4\frac{(\tau-\tau_0)\sigma^2}{1-\beta_2^2}\right| + \sup_{\tau_0<\tau\le\bar{\tau}}\frac{\left(\sum_{[\tau T]+1}^{T}\varepsilon_t y_{t-1}\right)^2}{T\sum_{[\tau T]+1}^{T} y_{t-1}^2}.$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

$$
\begin{aligned}
&\leq \left| \frac{1}{T} \sum_1^T \varepsilon_t^2 - \sigma^2 \right| + \left| \frac{2}{T} \sum_1^{[\tau_0 T]} y_{t-1}\varepsilon_t \right| \sup_{\tau_0<\tau\leq\bar{\tau}} |\hat{\beta}_1(\tau) - \beta_1| \\
&\quad + \left| \frac{1}{T} \sum_1^{[\tau_0 T]} y_{t-1}^2 - \frac{\tau_0\sigma^2}{1-\beta_1^2} \right| \sup_{\tau_0<\tau\leq\bar{\tau}} (\hat{\beta}_1(\tau) - \beta_1)^2 \\
&\quad + \frac{\tau_0\sigma^2}{1-\beta_1^2} \sup_{\tau_0<\tau\leq\bar{\tau}} |(\hat{\beta}_1(\tau) - \beta_1)^2 - \lambda_3| \\
&\quad + 2 \sup_{\tau_0<\tau\leq\bar{\tau}} |\hat{\beta}_1(\tau) - \beta_2| \sup_{\tau_0<\tau\leq\bar{\tau}} \left| \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t \right| \\
&\quad + \sup_{\tau_0<\tau\leq\bar{\tau}} (\hat{\beta}_1(\tau) - \beta_2)^2 \sup_{\tau_0<\tau\leq\bar{\tau}} \left| \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 - \frac{(\tau - \tau_0)\sigma^2}{1-\beta_2^2} \right| \\
&\quad + \sup_{\tau_0<\tau\leq\bar{\tau}} \frac{(\tau - \tau_0)\sigma^2}{1-\beta_2^2} \sup_{\tau_0<\tau\leq\bar{\tau}} |(\hat{\beta}_1(\tau) - \beta_2)^2 - \lambda_4| \\
&\quad + \sup_{\tau_0<\tau\leq\bar{\tau}} |\hat{\beta}_2(\tau) - \beta_2| \sup_{\tau_0<\tau\leq\bar{\tau}} \left| \frac{1}{T} \sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t \right| \\
&= o_p(1).
\end{aligned}
$$

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 $M_T \to \infty$, $(M_T/T) \to 0$ as $T \to \infty$, and

$$\lim_{T\to\infty} \Pr(|\hat{k} - k_0| > M_T) = \alpha,$$

where α is a positive constant in $(0,1]$.

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Note that

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

Because $k_0 \in D_{3T}$, we have $\inf_{m\in D_{3T}} RSS_T(m/T) \le RSS_T(\tau_0)$, and the preceding probability is bounded by

$$\le \Pr\left(\inf_{m\in D_{1T}\cup D_{2T}} RSS_T\left(\frac{m}{T}\right) < RSS_T(\tau_0)\right)$$

$$\le \sum_{i=1}^{2} \Pr\left(\inf_{m\in D_{iT}}\left[RSS_T\left(\frac{m}{T}\right) - RSS_T(\tau_0)\right] < 0\right).$$

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

$$= \Pr\left(\inf_{m\in D_{1T}}\left\{2\lambda\left(\frac{\sum_{k_0+1}^{T} y_{t-1}\varepsilon_t}{\sum_{k_0+1}^{T} y_{t-1}^2} - \frac{\sum_{m+1}^{k_0} y_{t-1}\varepsilon_t}{\sum_{m+1}^{k_0} y_{t-1}^2}\right)\right.\right.$$

$$\left.\left. + \lambda^2 + \frac{\sum_{m+1}^{T} y_{t-1}^2}{\sum_{k_0+1}^{T} y_{t-1}^2 \sum_{m+1}^{k_0} y_{t-1}^2}\Lambda_T\left(\frac{m}{T}\right)\right\} < 0\right)$$

$$+ \Pr\left(\inf_{m\in D_{2T}}\left\{2\lambda\left(\frac{\sum_{k_0+1}^{m} y_{t-1}\varepsilon_t}{\sum_{k_0+1}^{m} y_{t-1}^2} - \frac{\sum_{1}^{k_0} y_{t-1}\varepsilon_t}{\sum_{1}^{k_0} y_{t-1}^2}\right)\right.\right.$$

$$\left.\left. + \lambda^2 + \frac{\sum^{m} y_{t-1}^2}{\sum_{k_0+1}^{m} y_{t-1}^2 \sum_{1}^{k_0} y_{t-1}^2}\Lambda_T\left(\frac{m}{T}\right)\right\} < 0\right).$$

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

$$\le \Pr(-2\lambda A_1 + 2\lambda A_2 + A_3 > \lambda^2) + \Pr(2\lambda A_4 - 2\lambda A_5 + A_6 > \lambda^2),$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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;

$$A_2 = \sup_{m\in D_{1T}} \frac{\sum_{m+1}^{k_0} y_{t-1}\varepsilon_t}{\sum_{m+1}^{k_0} y_{t-1}^2} \le \sup_{m\in D_{1T}} \left| \frac{1}{k_0 - m} \sum_{m+1}^{k_0} \frac{y_{t-1}\varepsilon_t}{\sum_{m+1}^{k_0} y_{t-1}^2/(k_0 - m)} \right| = o_p(1)$$

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

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

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

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

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

$$A_4 = \frac{\sum_{1}^{k_0} y_{t-1}\varepsilon_t}{\sum_{1}^{k_0} y_{t-1}^2} = o_p(1)$$

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

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

by a similar argument as in A_2;

$$A_6 = \sup_{m\in D_{2T}}\left|\frac{\sum_{1}^{m} y_{t-1}^2}{\sum_{k_0+1}^{m} y_{t-1}^2 \sum_{1}^{k_0} y_{t-1}^2}\Lambda_T\left(\frac{m}{T}\right)\right| \le \left(\frac{1}{\sum_{1}^{k_0} y_{t-1}^2} + \frac{1}{\sum_{k_0+1}^{k_0+M_T} y_{t-1}^2}\right)$$

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

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

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

Because A_1–A_6 are $o_p(1)$ and λ^2 is a positive constant, we have

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

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

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

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

$$+ 1\{\hat{\tau}_T > \tau_0\}\sqrt{T}\left(\frac{\sum_{1}^{[\tau_0 T]} (\beta_1 y_{t-1} + \varepsilon_t) y_{t-1}}{\sum_{1}^{[\hat{\tau}_T T]} y_{t-1}^2} + \frac{\sum_{[\tau_0 T]+1}^{[\hat{\tau}_T T]} (\beta_2 y_{t-1} + \varepsilon_t) y_{t-1}}{\sum_{1}^{[\hat{\tau}_T T]} y_{t-1}^2}\right.$$

$$\left. - \frac{\sum_{1}^{[\tau_0 T]} (\beta_1 y_{t-1} + \varepsilon_t) y_{t-1}}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2}\right)$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

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

$$= 1\{\hat{\tau}_T \le \tau_0\}\sqrt{T}\left(O_p\left(\frac{1}{T}\right)O_p\left(\frac{1}{\sqrt{T}}\right) - O_p\left(\frac{1}{T}\right)\right)$$

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

$$= o_p(1).$$

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.

Define $\Im_t = \sigma(\varepsilon_i : i \le t)$ as the sigma field generated by the past history of $\{\varepsilon_t\}$. Because $\{\varepsilon_t y_{t-1}, \Im_t\}_{t=1}^{[\tau_0 T]}$ and $\{\varepsilon_t y_{t-1}, \Im_t\}_{t=[\tau_0 T]+1}^{T}$ are martingale difference sequences with

$$E(\varepsilon_t y_{t-1} | \Im_{t-1}) = 0 \qquad (t = 1,2,\ldots,T),$$

$$\frac{1}{[\tau_0 T]}\sum_{t=1}^{[\tau_0 T]} E((\varepsilon_t y_{t-1})^2 | \Im_{t-1}) \xrightarrow{p} \frac{\sigma^4}{1-\beta_1^2} < \infty,$$

$$\frac{1}{T-[\tau_0 T]}\sum_{t=[\tau_0 T]+1}^{T} E((\varepsilon_t y_{t-1})^2 | \Im_{t-1}) \xrightarrow{p} \frac{\sigma^4}{1-\beta_2^2} < \infty.$$

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

$$\sqrt{T}(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \overset{d}{=} \sqrt{T}(\hat{\beta}_1(\tau_0) - \beta_1) = \frac{\sum_{1}^{[\tau_0 T]} \varepsilon_t y_{t-1}/\sqrt{T}}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2/T} \xrightarrow{d} N\left(0, \frac{1-\beta_1^2}{\tau_0}\right),$$

$$\sqrt{T}(\hat{\beta}_2(\hat{\tau}_T) - \beta_2) \overset{d}{=} \sqrt{T}(\hat{\beta}_2(\tau_0) - \beta_1) = \frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}/\sqrt{T}}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2/T} \xrightarrow{d} N\left(0, \frac{1-\beta_2^2}{1-\tau_0}\right).$$

This proves Theorem 1. ■

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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 $\tau = \tau_0 + \upsilon[g(T)/T]$ and $\upsilon \le 0$, by (B.2) in Appendix B and the facts that

$$\text{(i)}\ \ \Lambda_T(\tau) = \frac{\left(\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t\right)^2}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2}\left(1 - \left(\frac{\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t}{\sum_{1}^{[\tau_0 T]} y_{t-1}\varepsilon_t}\right)^2 \frac{\sum_{1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{1}^{[\tau T]} y_{t-1}^2}\right)$$

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

$$= O_p(1)(1 - (1 - o_p(1))^2(1 + o_p(1)))$$

$$+ O_p(1)(1 - (1 + o_p(1))^2(1 - o_p(1))) = o_p(1);$$

$$\text{(ii)}\ \ (\beta_{2T} - \beta_1)\frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 \sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} = \frac{O_p(g(T))O_p(\sqrt{T})}{\sqrt{g(T)}O_p(T)} = O_p\left(\sqrt{\frac{g(T)}{T}}\right)$$

$$= o_p(1);$$

$$\text{(iii)}\ \ \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} \xrightarrow{p} 1,$$

we have

$$RSS_T(\tau) - RSS_T(\tau_0)$$

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

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

$$\Rightarrow -2\sigma^2\left(B_1\left(\frac{|\upsilon|}{1-\beta_1^2}\right) - \frac{1}{2}\frac{|\upsilon|}{(1-\beta_1^2)}\right),$$

where $B_1(\cdot)$ is a Brownian motion defined on R_+.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Similarly, for $\tau = \tau_0 + v[g(T)/T]$ and $v > 0$, from (B.4) in Appendix B and the fact that $[v/(1-\beta_{2T}^2)] \to [v/(1-\beta_1^2)]$ as $T \to \infty$, we have

$$
\begin{aligned}
&RSS_T(\tau) - RSS_T(\tau_0) \\
&\quad = 2(\beta_{2T} - \beta_1) \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t + (\beta_{2T} - \beta_1)^2 \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 + o_p(1) \\
&\quad = -\frac{2}{\sqrt{g(T)}} \sum_{0}^{[vg(T)]-1} (-y_{k_0+t}\varepsilon_{k_0+t+1}) + \frac{1}{g(T)} \sum_{0}^{[vg(T)]-1} y_{k_0+t}^2 + o_p(1) \\
&\quad \Rightarrow -2\sigma^2 \left(B_2\left(\frac{v}{1-\beta_1^2}\right) - \frac{v}{2(1-\beta_1^2)} \right),
\end{aligned}
$$

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

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

$$
\begin{aligned}
\frac{T(\beta_{2T} - \beta_1)^2}{1-\beta_1^2}(\hat{\tau}_T - \tau_0) = \hat{r} &= \operatorname*{Arg\,min}_r \{RSS_T(\tau) - RSS_T(\tau_0)\} \\
&\xrightarrow{d} \operatorname*{Arg\,min}_r \left\{ -2\sigma^2 \left(B^*(r) - \frac{1}{2}|r| \right) \right\} \\
&= \operatorname*{Arg\,max}_r \left(B^*(r) - \frac{1}{2}|r| \right),
\end{aligned}
$$

where $B^*(r)$ is a two-sided Brownian motion on R defined to be $B^*(r) = B_1(-r)$ for $r \le 0$ 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)
$$
\begin{aligned}
\frac{1}{T}\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t &= \frac{1}{2T}\sum_{[\tau_0 T]+1}^{T} (y_t^2 - y_{t-1}^2 - \varepsilon_t^2) \\
&= \frac{T-k_0}{2T}\left(\frac{y_T^2}{T-k_0} - \frac{y_{k_0}^2}{T-k_0} - \frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t^2}{T-k_0} \right) \\
&= \frac{1-\tau_0}{2}\left(\left(\frac{y_{k_0}}{\sqrt{T-k_0}} + \frac{1}{\sqrt{T-k_0}} \sum_{j=k_0+1}^{T} \varepsilon_j \right)^2 \right. \\
&\qquad \left. - \frac{y_{k_0}^2}{T-k_0} - \frac{1}{T-k_0} \sum_{[\tau_0 T]+1}^{T} \varepsilon_t^2 \right).
\end{aligned}
$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Because

$$\frac{y_{k_0}}{\sqrt{T-k_0}} \xrightarrow{p} 0 \quad \text{and} \quad \frac{1}{\sqrt{T-k_0}} \sum_{j=k_0+1}^{T} \varepsilon_j \Rightarrow \sigma B(1),$$

by the continuous mapping theorem, we have

$$\left(\frac{y_{k_0}}{\sqrt{T-k_0}} + \frac{1}{\sqrt{T-k_0}} \sum_{j=k_0+1}^{T} \varepsilon_j \right)^2 \Rightarrow \sigma^2 B^2(1).$$

Therefore

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

$$\text{(4c)} \quad \frac{1}{T^2} \sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 = \frac{(T-k_0)^2}{T^2} \sum_{[\tau_0 T]+1}^{T} \left(\frac{y_{k_0}}{\sqrt{T-k_0}} + \frac{1}{\sqrt{T-k_0}} \sum_{j=k_{0+1}}^{t-1} \varepsilon_j \right)^2 \frac{1}{T-k_0}$$

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

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

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

To prove (4g), we use the triangle inequality, (4b)–(4e), and the facts that

$$\sup_{\underline{\tau} \le \tau \le \tau_0} \sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 \le \sum_{1}^{[\tau_0 T]} y_{t-1}^2.$$

Then we have

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

$$\le \frac{1}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2} \left(|\beta_1 - 1| \sum_{1}^{[\tau_0 T]} y_{t-1}^2 + \sup_{\underline{\tau} \le \tau \le \tau_0} \left| \sum_{[\tau T]+1}^{[\tau_0 T]} \varepsilon_t y_{t-1} \right| \right.$$

$$\left. + \left| \sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1} \right| \right)$$

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

(4h) By the triangle inequality and (4g),

$$\sup_{\underline{\tau}\le\tau\le\tau_0} |\hat{\beta}_2(\tau)-\beta_1| \le \sup_{\underline{\tau}\le\tau\le\tau_0} |\hat{\beta}_2(\tau)-1| + |1-\beta_1| = O_p(1).$$

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

(4k) For $\tau \in \Xi_+^T$,

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

$$\le |\beta_1-1| \frac{\sum_{1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2} + \left| \frac{\sum_{1}^{[\tau_0 T]} \varepsilon_t y_{t-1}}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2} \right| + \left| \frac{\sum_{[\tau_0 T]+1}^{[\tau T]} \varepsilon_t y_{t-1}}{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2} \right|$$

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

$$+ \frac{T}{[(\tau-\tau_0)T]^2} A_T \left| \sum_{[\tau_0 T]+1}^{[\tau T]} \frac{y_{t-1}\varepsilon_t}{T} \right|,$$

where

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

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

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

(ii) By the definition of Ξ_+^T and by using the continuous mapping theorem, we have

$$A_T \Rightarrow \frac{1}{\sigma^2 \int_0^1 B^2} = O_p(1).$$

(4l) By the fact that

$$\sup_{\tau\in\Xi_+^T} \left| \frac{T^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} \right| \le \left| \frac{T^2}{\sum_{k+1}^{T} y_{t-1}^2} \right| = O_p(1)$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

and by (4j),

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

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

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

$$\sup_{\underline{\tau}\leq\tau\leq\tau_0}\left|\frac{1}{T}RSS_T(\tau)-\sigma^2-\frac{(1-\beta_1)(\tau_0-\tau)\sigma^2}{1+\beta_1}\right|$$

$$\leq \left|\frac{1}{T}\sum_{1}^{T}\varepsilon_t^2-\sigma^2\right|$$

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

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

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

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

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

$$=o_p(1).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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 $\sup|a-b| = \sup|a-c+c-b| \le \sup|a-c| + \sup|c-b|$, with

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

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

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 $\tau \in \Xi_+^T$, by (B.3) in Appendix B and the triangle inequality, we have

$$\begin{aligned}
&\left| \frac{1}{T} RSS_T(\tau) - \sigma^2 - \frac{(1-\beta_1)\tau_0\sigma^2}{1+\beta_1} \right| \\
&\quad \le \left| \frac{1}{T}\sum_1^T \varepsilon_t^2 - \sigma^2 \right| + \left| 2(\hat{\beta}_1(\tau) - \beta_1) \frac{1}{T} \sum_1^{[\tau_0 T]} y_{t-1}\varepsilon_t \right| \\
&\qquad + \left| (\hat{\beta}_1(\tau) - \beta_1)^2 \frac{\sum_1^{[\tau_0 T]} y_{t-1}^2}{T} - \frac{(1-\beta_1)\tau_0\sigma^2}{1+\beta_1} \right| + \left| 2(\hat{\beta}_1(\tau) - 1) \frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t}{T} \right| \\
&\qquad + (\hat{\beta}_1(\tau) - 1)^2 \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 + \left| (\hat{\beta}_2(\tau) - 1) \frac{1}{T} \sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t \right|.
\end{aligned}$$

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 $|a-b| = |a-c+c-b| \le |a-c| + |c-b|$, with

$$a = (\hat{\beta}_1(\tau) - \beta_1)^2 \frac{\sum_1^{[\tau_0 T]} y_{t-1}^2}{T}, \quad b = \frac{(1-\beta_1)\tau_0\sigma^2}{1+\beta_1} \quad \text{and} \quad c = (\hat{\beta}_1(\tau) - \beta_1)^2 \frac{\tau_0\sigma^2}{1-\beta_1^2}.$$

Thus, the preceding equation is bounded by

$$\begin{aligned}
&\le \left| \frac{1}{T}\sum_1^T \varepsilon_t^2 - \sigma^2 \right| + 2\left| \frac{1}{T} \sum_1^{[\tau_0 T]} y_{t-1}\varepsilon_t \right| |\hat{\beta}_1(\tau) - \beta_1| \\
&\quad + (\hat{\beta}_1(\tau) - \beta_1)^2 \left| \frac{1}{T} \sum_1^{[\tau_0 T]} y_{t-1}^2 - \frac{\tau_0\sigma^2}{1-\beta_1^2} \right| \\
&\quad + \frac{\tau_0\sigma^2}{1-\beta_1^2} |\hat{\beta}_1(\tau) + 1 - 2\beta_1||\hat{\beta}_1(\tau) - 1| + 2|\hat{\beta}_1(\tau) - 1| \left| \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t \right| \\
&\quad + \frac{[(\tau - \tau_0)T]^2}{T} (\hat{\beta}_1(\tau) - 1)^2 \frac{1}{[(\tau - \tau_0)T]^2} \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 \\
&\quad + |\hat{\beta}_2(\tau) - 1| \left| \frac{1}{T} \sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t \right|.
\end{aligned}$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

All seven preceding terms above are $o_p(1)$ by Lemma 2 and by the definition of Ξ_+^T. This derives equation (13).

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

$$\hat{\beta}_1\left(\tau_0 + \frac{c}{\sqrt{T}}\right) = \theta_T\left(\frac{1}{2}, c\right)\left(\beta_1 + \frac{\sum_1^{k_0} \varepsilon_t y_{t-1}}{\sum_1^{k_0} y_{t-1}^2}\right) + \left(1 - \theta_T\left(\frac{1}{2}, c\right)\right)\left(1 + \frac{\sum_{k_0+1}^{[k_0+c\sqrt{T}]} \varepsilon_t y_{t-1}}{\sum_{k_0+1}^{k_0+c\sqrt{T}} y_{t-1}^2}\right),$$

where

$$\theta_T\left(\frac{1}{2}, c\right) = \frac{\sum_1^{k_0} y_{t-1}^2}{\sum_1^{k_0} y_{t-1}^2 + \sum_{k_0+1}^{[k_0+c\sqrt{T}]} y_{t-1}^2}.$$

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

$$\frac{1}{T}\sum_{k_0+1}^{[k_0+c\sqrt{T}]} y_{t-1}^2 = \frac{1}{T}\sum_1^{[c\sqrt{T}]} y_{t+k_0-1}^2 \Rightarrow \sigma^2 \int_0^c B_3^2,$$

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

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

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

Additionally, we have

(iii) $\sup_{c\in R_{++}} \left|\hat{\beta}_2\left(\tau_0 + \dfrac{c}{\sqrt{T}}\right) - 1\right| = o_p(1)$;

(iv) $\sup_{c\in R_{++}} \dfrac{1}{T}\left|\sum_{k_0+1}^{[k_0+c\sqrt{T}]} y_{t-1}\varepsilon_t\right| = o_p(1)$;

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

$$\begin{aligned}
\frac{1}{T}RSS_T\left(\tau_0+\frac{c}{\sqrt{T}}\right) &= \frac{1}{T}\sum_1^T \varepsilon_t^2 - 2\left(\hat{\beta}_1\left(\tau_0+\frac{c}{\sqrt{T}}\right)-\beta_1\right)\frac{1}{T}\sum_1^{k_0} y_{t-1}\varepsilon_t \\
&\quad + \left(\hat{\beta}_1\left(\tau_0+\frac{c}{\sqrt{T}}\right)-\beta_1\right)^2\frac{1}{T}\sum_1^{k_0} y_{t-1}^2 \\
&\quad - 2\left(\hat{\beta}_1\left(\tau_0+\frac{c}{\sqrt{T}}\right)-1\right)\frac{1}{T}\sum_{k_0+1}^{[k_0+c\sqrt{T}]} y_{t-1}\varepsilon_t \\
&\quad + \left(\hat{\beta}_1\left(\tau_0+\frac{c}{\sqrt{T}}\right)-1\right)^2\frac{1}{T}\sum_{k_0+1}^{[k_0+c\sqrt{T}]} y_{t-1} \\
&\quad - \left(\hat{\beta}_2\left(\tau_0+\frac{c}{\sqrt{T}}\right)-1\right)\frac{1}{T}\sum_{[k_0+c\sqrt{T}]+1}^{T} y_{t-1}\varepsilon_t \\
&= \frac{1}{T}\sum_1^T \varepsilon_t^2 + o_p(1) + \left(\hat{\beta}_1\left(\tau_0+\frac{c}{\sqrt{T}}\right)-\beta_1\right)^2\frac{1}{T}\sum_1^{k_0} y_{t-1}^2 + o_p(1) \\
&\quad + \left(\hat{\beta}_1\left(\tau_0+\frac{c}{\sqrt{T}}\right)-1\right)^2\frac{1}{T}\sum_{k_0+1}^{[k_0+c\sqrt{T}]} y_{t-1}^2 + o_p(1) \\
&\Rightarrow \sigma^2 + \left(\frac{(1-\beta_1)(1-\beta_1^2)\int_0^c B_3^2}{\tau_0+(1-\beta_1^2)\int_0^c B_3^2}\right)^2\frac{\tau_0\sigma^2}{1-\beta_1^2} \\
&\quad + \left(\frac{\tau_0(\beta_1-1)}{\tau_0+(1-\beta_1^2)\int_0^c B_3^2}\right)^2\sigma^2\int_0^c B_3^2 \\
&= \sigma^2 + \frac{\tau_0\sigma^2(1-\beta_1)^2}{\tau_0\left(\int_0^c B_3^2\right)^{-1}+1-\beta_1^2}.
\end{aligned}$$

This derives equation (14).

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

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

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

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

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

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

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

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$A_4 = \frac{\sum_{1}^{k_0} y_{t-1}\varepsilon_t}{\sum_{1}^{k_0} y_{t-1}^2} = o_p(1)$$

by (4a) and (4e) of Lemma 2;

$$A_5 = \sup_{m\in D_{2T}} \frac{\sum_{k_0+1}^{m} y_{t-1}\varepsilon_t}{\sum_{k_0+1}^{m} y_{t-1}^2} = \sup_{m\in D_{2T}} \frac{y_m^2 - y_{k_0}^2 - \sum_{k_0+1}^{m} \varepsilon_t^2}{2\sum_{k_0+1}^{m} y_{t-1}^2} = o_p(1).$$

$$A_6 = \sup_{m\in D_{2T}} \left| \frac{\sum_{1}^{m} y_{t-1}^2}{\sum_{k_0+1}^{m} y_{t-1}^2 \sum_{1}^{k_0} y_{t-1}^2} \Lambda_T\left(\frac{m}{T}\right) \right| \le \left(\frac{1}{\sum_{1}^{k_0} y_{t-1}^2} + \frac{1}{\sum_{k_0+1}^{k_0+M_T} y_{t-1}^2} \right)$$

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

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

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

(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 $\sup_{m\in D_{2T}}(\sum_{k_0+1}^{m} y_{t-1}\varepsilon_t)/(\sum_{k_0+1}^{m} y_{t-1}^2)$ under nonstationarity of y_t. No study on the uniform convergence under this kind of nonstationarity has been done in the literature. However, the referees and I believe that the result is true and intuitive.)

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

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

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

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

$$= 1\{\hat{\tau}_T \le \tau_0\}\sqrt{T}\left(O_p\left(\frac{1}{T}\right)O_p\left(\frac{1}{\sqrt{T}}\right) - O_p\left(\frac{1}{T}\right)\right)$$

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

$$= o_p(1).$$

Thus, $\hat{\beta}_1(\hat{\tau}_T)$ and $\hat{\beta}_1(\tau_0)$ have the same asymptotic distribution.

Because $\{\varepsilon_t y_{t-1}, \Im_t\}_{t=1}^{[\tau_0 T]}$ is a martingale difference sequence, with

$$E(\varepsilon_t y_{t-1} | \Im_{t-1}) = 0,$$

$$\sum_{1}^{[\tau_0 T]} E((\varepsilon_t y_{t-1})^2 | \Im_{t-1}) \xrightarrow{p} \frac{\sigma^4}{1-\beta_1^2} < \infty.$$

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

$$\frac{1}{T}\sum_{1}^{[\tau_0 T]} y_{t-1}^2 \xrightarrow{p} \frac{\tau_0 \sigma^2}{1-\beta_1^2},$$

we have

$$\sqrt{T}(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \stackrel{d}{=} \sqrt{T}(\hat{\beta}_1(\tau_0) - \beta_1) = \frac{\sum_{1}^{[\tau_0 T]} \varepsilon_t y_{t-1}/\sqrt{T}}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2/T} \xrightarrow{d} N\left(0, \frac{1-\beta_1^2}{\tau_0}\right).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

$$= 1\{\hat{\tau}_T \le \tau_0\} T \left(-\frac{\sum_{[\hat{\tau}_T T]+1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{[\hat{\tau}_T T]+1}^{T} y_{t-1}^2} \frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2} + \frac{\sum_{[\hat{\tau}_T T]+1}^{[\tau_0 T]} \varepsilon_t y_{t-1}}{\sum_{[\hat{\tau}_T T]+1}^{T} y_{t-1}^2} \right.$$

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

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

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

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

$$= o_p(1).$$

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

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

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 β_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 υ be a finite constant, and $B_1(\cdot)$ and $B_2(\cdot)$ be defined as in Section 2. For $\tau = \tau_0 + \upsilon[g(T)/T]$ and $\upsilon \le 0$, we have $\Lambda_T(\tau) = o_p(1)$ where $\Lambda_T(\tau)$ is defined in Appendix B, and

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$\frac{y_{k_0-t-1}}{\sqrt{g(T)}} = \frac{1}{\sqrt{g(T)}}\left[\left(1-\frac{1}{g(T)}\right)^{k_0-t-2} y_0 + \sum_{i=0}^{k_0-t-2}\left(1-\frac{1}{g(T)}\right)^{i}\varepsilon_{k_0-t-1-i}\right]$$

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

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

where $B_a(\frac{1}{2})$ is a Brownian motion defined on R_+.

Because

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

we have

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

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

Further, because

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

and

$$\frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} \xrightarrow{p} 1,$$

equation (B.2) in Appendix B becomes

$$RSS_T(\tau) - RSS_T(\tau_0)$$

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

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

$$\Rightarrow -2\sigma^2 B_a\left(\frac{1}{2}\right)B_1(|v|) + |v|\sigma^2 B_a^2\left(\frac{1}{2}\right).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Similarly, for $\tau = \tau_0 + \upsilon[g(T)/T]$ and $\upsilon > 0$, from (B.4) in Appendix B, we have

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

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

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

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

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

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

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

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

$$\begin{aligned}
&RSS_T(\tau) - RSS_T(\tau_0)\\
&\quad = 2(1-\beta_{1T})\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t + (1-\beta_{1T})^2\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 + o_p(1)\\
&\quad = \frac{2}{g(T)}\sum_{0}^{[\upsilon g(T)]-1} y_{k_0+t}\varepsilon_{k_0+t+1} + \frac{1}{g^2(T)}\sum_{0}^{[\upsilon g(T)]-1} y_{k_0+t}^2 + o_p(1)\\
&\quad \Rightarrow 2\sigma^2\int_0^{\upsilon}\left(B_2(s) + B_a\left(\frac{1}{2}\right)\right)dB_2(s) + \sigma^2\int_0^{\upsilon}\left(B_2(s) + B\left(\frac{1}{2}\right)\right)^2 ds\\
&\quad = -2\sigma^2 B_a^2\left(\frac{1}{2}\right)\\
&\qquad \times\left[-\frac{B_2(\upsilon)}{B_a\left(\frac{1}{2}\right)} - \int_0^{\upsilon}\frac{B_2(s)}{B_a^2\left(\frac{1}{2}\right)}dB_2(s) - \int_0^{\upsilon}\left(\frac{B_2(s)}{2B_a\left(\frac{1}{2}\right)} + 1\right)\frac{B_2(s)}{B_a\left(\frac{1}{2}\right)}ds - \frac{1}{2}\upsilon\right].
\end{aligned}$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

By the continuous mapping theorem for argmax functionals, we have

$$(1-\beta_{1T})T(\hat{\tau}_T-\tau_0)=\hat{\upsilon}=\underset{\upsilon\in R}{\text{Arg max}}\{RSS_T(\tau)-RSS_T(\tau_0)\}$$

$$\xrightarrow{d}\underset{\upsilon\in R}{\text{Arg max}}\left\{-2\sigma^2B_a^2\left(\frac{1}{2}\right)\left(\frac{C^*(\upsilon)}{B_a\left(\frac{1}{2}\right)}-\frac{1}{2}|\upsilon|\right)\right\}$$

$$=\underset{\upsilon\in R}{\text{Arg max}}\left(\frac{C^*(\upsilon)}{B_a\left(\frac{1}{2}\right)}-\frac{1}{2}|\upsilon|\right).$$

This proves Theorem 3. ■

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

$$\frac{1}{T}\sum_1^{[\tau_0T]}y_{t-1}\varepsilon_t=\frac{1}{2T}\sum_1^{[\tau_0T]}(y_t^2-y_{t-1}^2-\varepsilon_t^2)=\frac{1}{2T}\left(y_{k_0}^2-y_0^2-\sum_1^{[\tau_0T]}\varepsilon_t^2\right)$$

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

(5b)
$$\frac{1}{\sqrt{T}}\sum_{[\tau_0T]+1}^{T}y_{t-1}\varepsilon_t$$

$$=\frac{1}{\sqrt{T}}\sum_{t=k_0+1}^{T}\left(\beta_2^{t-k_0-1}y_{k_0}+\sum_{i=k_0+1}^{t-1}\beta_2^{t-i-1}\varepsilon_i\right)\varepsilon_t$$

$$=\frac{1}{\sqrt{T}}\left(\left(y_0+\sum_{j=1}^{k_0}\varepsilon_j\right)\sum_{k_0+1}^{T}\beta_2^{t-k_0-1}\varepsilon_t+\sum_{t=k_0+1}^{T}\sum_{i=k_0+1}^{t-1}\beta_2^{t-i-1}\varepsilon_i\varepsilon_t\right)$$

$$=\frac{1}{\sqrt{T}}\sum_{j=1}^{k_0}\left(\varepsilon_j\sum_{k_0+1}^{T}\beta_2^{t-k_0-1}\varepsilon_t\right)+\frac{1}{\sqrt{T}}\sum_{t=k_0+1}^{T}\left(\varepsilon_t\sum_{i=k_0+1}^{t-1}\beta_2^{t-i-1}\varepsilon_i\right)$$

$$+o_p(1).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

By Assumption (A2) that ε's are i.i.d., $\{\varepsilon_j \sum_{t=k_0+1}^{T} \beta_2^{t-k_0-1}\varepsilon_t, \Im_j\}_{j=1}^{k_0}$ and $\{\varepsilon_t \sum_{i=k_0+1}^{t-1} \beta_2^{t-i-1}\varepsilon_i, \Im_t\}_{t=k_0+1}^{T}$ will both be martingale difference sequences with

$$E\left(\varepsilon_j \sum_{t=k_0+1}^{T} \beta_2^{t-k_0-1}\varepsilon_t | \Im_{j-1}\right) = 0 \qquad (j = 1,2,\ldots,k_0),$$

$$\frac{1}{T}\sum_{j=1}^{k_0} E\left(\left(\varepsilon_j \sum_{t=k_0+1}^{T} \beta_2^{t-k_0-1}\varepsilon_t\right)^2 | \Im_{j-1}\right) \xrightarrow{p} \frac{\tau_0\sigma^4}{1-\beta_2^2} < \infty,$$

$$E\left(\varepsilon_t \sum_{i=k_0+1}^{t-1} \beta_2^{t-i-1}\varepsilon_i | \Im_{t-1}\right) = 0 \qquad (t = k_0+1, k_0+2,\ldots,T),$$

$$\frac{1}{T}\sum_{t=k_0+1}^{T} E\left(\left(\varepsilon_t \sum_{i=k_0+1}^{t-1} \beta_2^{t-i-1}\varepsilon_i\right)^2 | \Im_{t-1}\right) \xrightarrow{p} \frac{(1-\tau_0)\sigma^4}{1-\beta_2^2}.$$

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

$$\frac{1}{\sqrt{T}}\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t \xrightarrow{d} N\left(0, \frac{\sigma^4}{1-\beta_2^2}\right);$$

$$\text{(5c)} \quad \frac{1}{T^2}\sum_{1}^{[\tau_0 T]} y_{t-1}^2 = \sum_{1}^{[\tau_0 T]}\left(\frac{y_0}{\sqrt{T}} + \frac{\sum_{i=1}^{t-1}\varepsilon_i}{\sqrt{T}}\right)^2 \frac{1}{T} \Rightarrow \sigma^2 \int_0^{\tau_0} B^2 = O_p(1);$$

$$\begin{aligned}\text{(5d)} \quad \frac{1}{T}\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 &= \frac{1}{T}\sum_{t=k_0+1}^{[\tau T]}\left(\beta_2^{t-k_0-1} y_{k_0} + \sum_{i=k_0+1}^{t-1}\beta_2^{t-i-1}\varepsilon_i\right)^2 \\ &= \left(\frac{y_{k_0}}{\sqrt{T}}\right)^2 \sum_{t=k_0+1}^{[\tau T]} \beta_2^{2(t-k_0-1)} \\ &\quad + 2\frac{y_{k_0}}{T}\sum_{t=k_0+1}^{[\tau T]}\left(\beta_2^{t-k_0-1}\sum_{i=k_0+1}^{t-1}\beta_2^{t-i-1}\varepsilon_i\right) \\ &\quad + \frac{1}{T}\sum_{t=k_0+1}^{[\tau T]}\left(\sum_{i=k_0+1}^{t-1}\beta_2^{t-i-1}\varepsilon_i\right)^2.\end{aligned}$$

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

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

$$\leq 2\left|\frac{y_{k_0}}{\sqrt{T}}\right| \sup_{\tau\in\Xi_+^\tau}\left|\sum_{k_0+1}^{[\tau T]}\beta_2^{t-k_0-1}\frac{1-\beta_2^{t-k_0-1}}{1-\beta_2}\right| \frac{1}{\sqrt{T}}\sup_{t>k_0}|\varepsilon_t|$$

$$\leq 2\left|\frac{y_{k_0}}{\sqrt{T}}\right|\left(\sum_{k_0+1}^{\infty}\left|\frac{\beta_2^{t-k_0-1}}{1-\beta_2}\right| + \sum_{k_0+1}^{\infty}\left|\frac{\beta_2^{2(t-k_0-1)}}{1-\beta_2}\right|\right)\frac{1}{\sqrt{T}}\sup_{t>k_0}|\varepsilon_t|$$

$$\leq 2\left|\frac{y_{k_0}}{\sqrt{T}}\right|\left(\frac{1}{(1-|\beta_2|)^2} + \frac{1}{(1-|\beta_2|)(1-\beta_2^2)}\right)\frac{1}{\sqrt{T}}\sup_{t>k_0}|\varepsilon_t| = o_p(1).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Last, because $|\beta_2| < 1$ and ε_i are i.i.d. with finite fourth moment, we have

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

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

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

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

Thus,

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

(5e) $$\sup_{\tau\in\Xi_-^T} \left| \frac{1}{T} \sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t \right| = \sup_{\tau\in\Xi_-^T} \left| \sum_{1}^{[\tau T]} \frac{y_{t-1}}{\sqrt{T}} \frac{\varepsilon_t}{\sqrt{T}} \right| \xrightarrow{d} \sup_{\tau\in\Xi_-^T} \left| \sigma^2 \int_0^\tau B dB \right| = O_p(1);$$

(5f) $$\sup_{\tau\in\Xi_-^T} \left| \frac{1}{T} \sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t \right| = \sup_{\tau\in\Xi_-^T} \left| \sum_{[\tau T]+1}^{[\tau_0 T]} \frac{y_{t-1}}{\sqrt{T}} \frac{\varepsilon_t}{\sqrt{T}} \right|$$

$$\xrightarrow{d} \sup_{\tau\in\Xi_-^T} \left| \sigma^2 \int_\tau^{\tau_0} B dB \right| = O_p(1);$$

(5g) $$\sup_{\tau\in\Xi_-^T} |\hat{\beta}_1(\tau) - 1| = \sup_{\tau\in\Xi_-^T} \left| \frac{\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t}{\sum_{1}^{[\tau T]} y_{t-1}^2} \right| \le \frac{1}{T} \frac{T^2}{\sum_{1}^{[\tau T]} y_{t-1}^2} \sup_{\tau\in\Xi_-^T} \left| \frac{\sum_{1}^{[\tau T]} y_{t-1}\varepsilon_t}{T} \right|$$

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

(5h) $$\sup_{\tau\in\Xi_-^T} |\hat{\beta}_2(\tau) - 1| = \sup_{\tau\in\Xi_-^T} \left| \frac{(\beta_2 - 1) \sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 + \sum_{[\tau T]+1}^{[\tau_0 T]} \varepsilon_t y_{t-1} + \sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 + \sum_{[\tau_0 T]+1}^{T} y_{t-1}^2} \right|$$

$$\le \sup_{\tau\in\Xi_-^T} \frac{|\beta_2 - 1| \sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2} + \sup_{\tau\in\Xi_-^T} \left| \frac{\sum_{[\tau T]+1}^{[\tau_0 T]} \varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2} \right|$$

$$+ \left| \frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2} \right|.$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Using the fact that

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

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

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

The second term is bounded by

$$\sup_{\tau \in \Xi_-^T} \frac{1}{k_0 - k} \sup_{\tau \in \Xi_-^T} \left| \frac{T(k_0 - k)}{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2} \right| \sup_{\tau \in \Xi_-^T} \left| \frac{\sum_{[\tau T]+1}^{[\tau_0 T]} \varepsilon_t y_{t-1}}{T} \right|$$

$$= O_p(T^{-1/2}) O_p(1) O_p(1) = o_p(1).$$

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

$$\text{(5i)} \quad \sup_{\tau \in \Xi_+^T} \left| \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1} \varepsilon_t \right| = \sup_{\tau \in \Xi_+^T} \left| \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} \left(\beta_2^{t-k_0-1} y_{k_0} + \sum_{i=k_0+1}^{t-1} \beta_2^{t-i-1} \varepsilon_i \right) \varepsilon_t \right|$$

$$\leq \sup_{\tau \in \Xi_+^T} \left| \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} \beta_2^{t-k_0-1} y_{k_0} \varepsilon_t \right|$$

$$+ \sup_{\tau \in \Xi_+^T} \left| \frac{1}{T} \sum_{[\tau_0 T]+1}^{[\tau T]} \sum_{i=k_0+1}^{t-1} \beta_2^{t-i-1} \varepsilon_i \varepsilon_t \right|.$$

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

Further, because $E(\sum_{i=k_0+1}^{t-1} \beta_2^{t-i-1} \varepsilon_i \varepsilon_t) = 0$ and $\text{Var}(\sum_{i=k_0+1}^{t-1} \beta_2^{t-i-1} \varepsilon_i \varepsilon_t) \leq \sigma^4/(1 - \beta_2^2) < \infty$, 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).

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

To show (5j), by (5b), (5i), and the triangle inequality, we have

$$\sup_{\tau\in\Xi_+^T}\left|\frac{1}{T}\sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t\right| \le \left|\frac{1}{T}\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t\right| + \sup_{\tau\in\Xi_+^T}\left|\frac{1}{T}\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t\right| = o_p(1).$$

(5k)
$$\begin{aligned}\sup_{\tau\in\Xi_+^T}|\hat{\beta}_1(\tau)-1| &= \sup_{\tau\in\Xi_+^T}\left|\frac{(\beta_2-1)\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 + \sum_{1}^{[\tau_0 T]}\varepsilon_t y_{t-1} + \sum_{[\tau_0 T]+1}^{[\tau T]}\varepsilon_t y_{t-1}}{\sum_{1}^{[\tau T]} y_{t-1}^2}\right| \\ &\le \frac{|\beta_2-1|\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2 + \left|\sum_{1}^{[\tau_0 T]}\varepsilon_t y_{t-1}\right| + \sup_{\tau\in\Xi_+^T}\left|\sum_{[\tau_0 T]+1}^{[\tau T]}\varepsilon_t y_{t-1}\right|}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2} \\ &= \frac{O_p(T)+O_p(T)+o_p(T)}{O_p(T^2)} = O_p\left(\frac{1}{T}\right)\end{aligned}$$

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

(5l)
$$\sup_{\tau\in\Xi_+^T}|\hat{\beta}_2(\tau)-\beta_2| = \sup_{\tau\in\Xi_+^T}\left|\frac{\sum_{[\tau T]+1}^{T}\varepsilon_t y_{t-1}}{\sum_{[\tau T]+1}^{T} y_{t-1}^2}\right| \le \frac{T}{\sum_{[\bar{\tau} T]+1}^{T} y_{t-1}^2}\sup_{\tau\in\Xi_+^T}\left|\frac{\sum_{[\tau T]+1}^{T}\varepsilon_t y_{t-1}}{T}\right|$$
$$= O_p(1)o_p(1) = o_p(1)$$

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 σ^2 for $\tau \in \Xi_-^T$, it is sufficient to show that

(i) $(1/T)RSS_T(\tau)$ converges pointwise to a random value above σ^2 for all $\tau \in \Xi_-^T$ and
(ii) $\sup_{\tau_b,\tau_a\in\Xi_-^T}|(1/T)RSS_T(\tau_b) - (1/T)RSS_T(\tau_a)| \xrightarrow{p} 0$ for any $\tau_b, \tau_a \in \Xi_-^T$.

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

To show (i), use Lemma 3 and (B.1) in Appendix B. For $\tau \in \Xi_-^T$, $\beta_1 = 1$, and $|\beta_2| < 1$, we have

$$\begin{aligned}
\frac{1}{T} RSS_T(\tau) = {} & \frac{\sum_1^T \varepsilon_t^2}{T} - (\hat{\beta}_1(\tau) - 1) \frac{\sum_1^{[\tau T]} y_{t-1}\varepsilon_t}{T} - 2(\hat{\beta}_2(\tau) - 1) \frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t}{T} \\
& + \frac{1}{T}(T(\hat{\beta}_2(\tau) - 1))^2 \frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2}{T^2} - \frac{2}{\sqrt{T}}(\hat{\beta}_2(\tau) - \beta_2) \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t}{\sqrt{T}} \\
& + (\hat{\beta}_2(\tau) - 1)^2 \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{T} + 2(\hat{\beta}_2(\tau) - 1)(1 - \beta_2) \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{T} \\
& + (1 - \beta_2)^2 \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{T} \\
= {} & \frac{\sum_1^T \varepsilon_t^2}{T} - o_p(1) - o_p(1) + o_p(1) - o_p(1) + o_p(1) + o_p(1) + (1 - \beta_2)^2 \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{T} \\
\Rightarrow {} & \sigma^2 + \frac{(1 - \beta_2)(1 - \tau_0 + B^2(\tau_0))}{1 + \beta_2} \sigma^2 > \sigma^2.
\end{aligned}$$

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

$$\begin{aligned}
& \sup_{\tau_b, \tau_a \in \Xi_-^T} \left| \frac{1}{T} RSS_T(\tau_b) - \frac{1}{T} RSS_T(\tau_a) \right| \\
= {} & \sup_{\tau_b, \tau_a \in \Xi_-^T} \Bigg| - (\hat{\beta}_1(\tau_b) - 1) \frac{\sum_1^{[\tau_b T]} y_{t-1}\varepsilon_t}{T} - 2(\hat{\beta}_2(\tau_b) - 1) \frac{\sum_{[\tau_b T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t}{T} \\
& + (\hat{\beta}_2(\tau_b) - 1)^2 \frac{\sum_{[\tau_b T]+1}^{[\tau_0 T]} y_{t-1}^2}{T} + 2(\hat{\beta}_2(\tau_a) - \hat{\beta}_2(\tau_b)) \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t}{T} \\
& + ((\hat{\beta}_2(\tau_b) - \beta_2)^2 - (\hat{\beta}_2(\tau_a) - \beta_2)^2) - \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{T} + (\hat{\beta}_1(\tau_a) - 1) \frac{\sum_1^{[\tau_a T]} y_{t-1}\varepsilon_t}{T} \\
& + 2(\hat{\beta}_2(\tau_a) - 1) \frac{\sum_{[\tau_a T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t}{T} - (\hat{\beta}_2(\tau_a) - 1)^2 \frac{\sum_{[\tau_a T]+1}^{[\tau_0 T]} y_{t-1}^2}{T} \Bigg|
\end{aligned}$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$\leq \sup_{\tau_b \in \Xi_-^T} |\hat{\beta}_1(\tau_b) - 1| \sup_{\tau_b \in \Xi_-^T} \left| \frac{\sum_{1}^{[\tau_b T]} y_{t-1}\varepsilon_t}{T} \right|$$

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

$$+ \sup_{\tau_b \in \Xi_-^T} (\hat{\beta}_2(\tau_b) - 1)^2 \sup_{\tau_b \in \Xi_-^T} \frac{\sum_{[\tau_b T]+1}^{[\tau_0 T]} y_{t-1}^2}{T} + 2 \left| \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t}{T} \right| \sup_{\tau_b, \tau_a \in \Xi_-^T} |\hat{\beta}_2(\tau_a) - \hat{\beta}_2(\tau_b)|$$

$$+ \frac{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2}{T} \sup_{\tau_b, \tau_a \in \Xi_-^T} |(\hat{\beta}_2(\tau_b) - \beta_2)^2 - (\hat{\beta}_2(\tau_a) - \beta_2)^2|$$

$$+ \sup_{\tau_a \in \Xi_-^T} |\hat{\beta}_1(\tau_a) - 1| \sup_{\tau_a \in \Xi_-^T} \left| \frac{\sum_{1}^{[\tau_a T]} y_{t-1}\varepsilon_t}{T} \right| + 2 \sup_{\tau_a \in \Xi_-^T} |\hat{\beta}_2(\tau_a) - 1| \sup_{\tau_a \in \Xi_-^T} \left| \frac{\sum_{[\tau_a T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t}{T} \right|$$

$$+ \sup_{\tau_a \in \Xi_-^T} (\hat{\beta}_2(\tau_a) - 1)^2 \sup_{\tau_a \in \Xi_-^T} \frac{\sum_{[\tau_a T]+1}^{[\tau_0 T]} y_{t-1}^2}{T} \xrightarrow{p} 0.$$

This derives equation (16).

For $\tau = \tau_0$,

$$\frac{1}{T} RSS_T(\tau_0)$$

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

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

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

This derives equation (17).

To show that $(1/T)RSS_T(\tau)$ converges uniformly to a random linear function above σ^2 for $\tau \in \Xi_+^T$, it is sufficient to show that

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

(i) $(1/T)RSS_T(\tau)$ converges pointwise to a random linear function above σ^2 for all $\tau \in \Xi_+^T$ and

(ii) $\sup_{\tau\in\Xi_+^T}|(1/T)RSS_T(\tau) - (1/T)RSS_T(\tau + (1/T))| \xrightarrow{p} 0$ for all $\tau \in \Xi_+^T$.

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

$$\begin{aligned}
\frac{1}{T}RSS_T(\tau) &= \frac{\sum_1^T \varepsilon_t^2}{T} - 2(\hat{\beta}_1(\tau)-1)\frac{\sum_1^{[\tau_0 T]} y_{t-1}\varepsilon_t}{T} + (\hat{\beta}_1(\tau)-1)^2\frac{\sum_1^{[\tau_0 T]} y_{t-1}^2}{T} \\
&\quad - 2(\hat{\beta}_1(\tau)-\beta_2)\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t}{T} + (\hat{\beta}_1(\tau)-\beta_2)^2\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2}{T} \\
&\quad - (\hat{\beta}_2(\tau)-\beta_2)\frac{\sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t}{T} \\
&= \frac{\sum_1^T \varepsilon_t^2}{T} - o_p(1) + o_p(1) - o_p(1) + (\hat{\beta}_1(\tau)-\beta_2)^2\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2}{T} - o_p(1) \\
&\Rightarrow \sigma^2 + \frac{(1-\beta_2)^2(\tau - \tau_0 + B^2(\tau_0))}{1-\beta_2^2}\sigma^2 \geq \sigma^2.
\end{aligned}$$

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

$$\begin{aligned}
&\sup_{\tau\in\Xi_+^T}\left|\frac{1}{T}RSS_T(\tau) - \frac{1}{T}RSS_T\left(\tau+\frac{1}{T}\right)\right| \\
&= \sup_{\tau\in\Xi_+^T}\Bigg| 2\left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\hat{\beta}_1(\tau)\right)\frac{\sum_1^{[\tau_0 T]} y_{t-1}\varepsilon_t}{T} \\
&\quad + \left((\hat{\beta}_1(\tau)-1)^2 - \left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-1\right)^2\right)\frac{\sum_1^{[\tau_0 T]} y_{t-1}^2}{T} \\
&\quad + 2\left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\hat{\beta}_1(\tau)\right)\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t}{T} + 2\left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\beta_2\right)\frac{y_{[\tau T]}\varepsilon_{[\tau T]+1}}{T} \\
&\quad + \left((\hat{\beta}_1(\tau)-\beta_2)^2 - \left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\beta_2\right)^2\right)\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2}{T} \\
&\quad - \left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\beta_2\right)^2\frac{y_{[\tau T]}^2}{T} + \left(\hat{\beta}_2\left(\tau+\frac{1}{T}\right)-\hat{\beta}_2(\tau)\right)\frac{\sum_{[\tau T]+1}^{T} y_{t-1}\varepsilon_t}{T} \\
&\quad - \left(\hat{\beta}_2\left(\tau+\frac{1}{T}\right)-\beta_2\right)\frac{y_{[\tau T]}\varepsilon_{[\tau T]+1}}{T}\Bigg|
\end{aligned}$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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

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

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

$$+2\sup_{\tau\in\Xi_+^T}\left|\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\beta_2\right|\sup_{\tau\in\Xi_+^T}\left|\frac{y_{[\tau T]}\varepsilon_{[\tau T]+1}}{T}\right|$$

$$+\sup_{\tau\in\Xi_+^T}\left|(\hat{\beta}_1(\tau)-\beta_2)^2-\left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\beta_2\right)^2\right|\sup_{\tau\in\Xi_+^T}\frac{\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2}{T}$$

$$+\sup_{\tau\in\Xi_+^T}\left(\hat{\beta}_1\left(\tau+\frac{1}{T}\right)-\beta_2\right)^2\sup_{\tau\in\Xi_+^T}\frac{y_{[\tau T]}^2}{T}$$

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

$$+\sup_{\tau\in\Xi_+^T}\left|\hat{\beta}_2\left(\tau+\frac{1}{T}\right)-\beta_2\right|\sup_{\tau\in\Xi_+^T}\left|\frac{y_{[\tau T]}\varepsilon_{[\tau T]+1}}{T}\right|$$

$$\xrightarrow{p} 0.$$

This derives equation (18).

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

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

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

(Jb) $$\sup_{m\in\{1,2,..\},(m/T)\to 0} \frac{1}{T}\left|\sum_{k_0-m+1}^{k_0} y_{t-1}\varepsilon_t\right| = o_p(1);$$

(Jc) $$\sup_{m\in\{1,2,..\},(m/T)\to 0} \left|\hat{\beta}_1\left(\tau_0-\frac{m}{T}\right)-1\right| = o_p(1);$$

$$\hat{\beta}_2\left(\tau_0-\frac{m}{T}\right) = \alpha\left(\tau_0-\frac{m}{T}\right)\left(1+\frac{\sum_{k_0-m+1}^{k_0}\varepsilon_t y_{t-1}}{\sum_{k_0-m+1}^{k_0} y_{t-1}^2}\right)$$

$$+\left(1-\alpha\left(\tau_0-\frac{m}{T}\right)\right)\left(\beta_2+\frac{\sum_{k_0+1}^{T}\varepsilon_t y_{t-1}}{\sum_{k_0+1}^{T} y_{t-1}^2}\right),$$

where

$$\alpha\left(\tau_0-\frac{m}{T}\right) = \frac{\frac{1}{T}\sum_{k_0-m+1}^{k_0} y_{t-1}^2}{\frac{1}{T}\sum_{k_0-m+1}^{k_0} y_{t-1}^2+\frac{1}{T}\sum_{k_0+1}^{T} y_{t-1}^2} \Rightarrow \frac{mB^2(\tau_0)}{mB^2(\tau_0)+\frac{1-\tau_0+B^2(\tau_0)}{1-\beta_2^2}}.$$

Further, because

$$\frac{\sum_{k_0-m+1}^{k_0}\varepsilon_t y_{t-1}}{\sum_{k_0-m+1}^{k_0} y_{t-1}^2} = o_p(1), \quad \frac{\sum_{k_0+1}^{T}\varepsilon_t y_{t-1}}{\sum_{k_0+1}^{T} y_{t-1}^2} = o_p(1),$$

we have

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

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).

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Using (B.1) in Appendix B, (5a), (5b), and (5d) in Lemma 3, and (Ja)–(Jd), we have

$$\begin{aligned}
&\frac{1}{T} RSS_T\left(\tau_0 - \frac{m}{T}\right) \\
&= \frac{\sum_1^T \varepsilon_t^2}{T} - \frac{\left(\sum_1^{k_0-m} y_{t-1}\varepsilon_t\right)^2}{T\sum_1^{k_0-m} y_{t-1}^2} - 2\left(\hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) - 1\right)\frac{\sum_{k_0-m+1}^{k_0} y_{t-1}\varepsilon_t}{T} \\
&\quad + \left(\hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) - 1\right)^2 \frac{\sum_{k_0-m+1}^{k_0} y_{t-1}^2}{T} \\
&\quad - 2\left(\hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) - \beta_2\right)\frac{\sum_{k_0+1}^{T} y_{t-1}\varepsilon_t}{T} + \left(\hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) - \beta_2\right)^2 \frac{\sum_{k_0+1}^{T} y_{t-1}^2}{T} \\
&= \frac{\sum_1^T \varepsilon_t^2}{T} + o_p(1) + o_p(1) + \left(\hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) - 1\right)^2 \frac{\sum_{k_0-m+1}^{k_0} y_{t-1}^2}{T} + o_p(1) \\
&\quad + \left(\hat{\beta}_2\left(\tau_0 - \frac{m}{T}\right) - \beta_2\right)^2 \frac{\sum_{k_0+1}^{T} y_{t-1}^2}{T} \\
&\Rightarrow \sigma^2 + \frac{(1-\tau_0+B^2(\tau_0))(1-\beta_2)^2 m\sigma^2 B^2(\tau_0)}{(1-\beta_2^2)mB^2(\tau_0) + 1 - \tau_0 + B^2(\tau_0)}.
\end{aligned}$$

This derives equation (20).

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

(Je) $$\sup_{m\in\{1,2,..\},(m/T)\to 0} \frac{1}{T}\left|\sum_{k_0+1}^{k_0+m} y_{t-1}\varepsilon_t\right| = o_p(1);$$

(Jf) $$\frac{1}{T}\sum_{k_0+1}^{k_0+m} y_{t-1}^2 \Rightarrow \frac{\sigma^2 B^2(\tau_0)(1-\beta_2^{2m})}{1-\beta_2^2}$$

by (5d) in Lemma 3;

(Jg) $$\sup_{m\in\{1,2,..\},(m/T)\to 0} \frac{1}{T}\left|\sum_{k_0+m+1}^{T} y_{t-1}\varepsilon_t\right| = o_p(1);$$

(Jh) $$\frac{1}{T}\sum_{k_0+m+1}^{T} y_{t-1}^2 \le \frac{1}{T}\sum_{k_0+1}^{T} y_{t-1}^2 = O_p(1)$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

by (5d) in Lemma 3;

$$\text{(Ji)} \quad \sup_{m\in\{1,2,..\},(m/T)\to 0} \left|\hat{\beta}_1\left(\tau_0+\frac{m}{T}\right)-1\right| = O_p\left(\frac{1}{T}\right);$$

$$\text{(Jj)} \quad \sup_{m\in\{1,2,..\},(m/T)\to 0} \left|\hat{\beta}_2\left(\tau_0+\frac{m}{T}\right)-\beta_2\right| = o_p(1).$$

Using (B.3) in Appendix B, by (5a)–(5c) of Lemma 3 and (Je) to (Jj) given previously, we can write the residual sum of squares as follows:

$$\begin{aligned}
\frac{1}{T}RSS_T\left(\tau_0+\frac{m}{T}\right) &= \frac{\sum_1^T \varepsilon_t^2}{T} - 2\left(\hat{\beta}_1\left(\tau_0+\frac{m}{T}\right)-1\right)\frac{\sum_1^{k_0} y_{t-1}\varepsilon_t}{T} \\
&\quad + \left(\hat{\beta}_1\left(\tau_0+\frac{m}{T}\right)-1\right)^2 \frac{\sum_1^{k_0} y_{t-1}^2}{T} \\
&\quad - 2\left(\hat{\beta}_1\left(\tau_0+\frac{m}{T}\right)-\beta_2\right)\frac{\sum_{k_0+1}^{k_0+m} y_{t-1}\varepsilon_t}{T} + \left(\hat{\beta}_1\left(\tau_0+\frac{m}{T}\right)-\beta_2\right)^2 \\
&\quad \times \frac{\sum_{k_0+1}^{k_0+m} y_{t-1}^2}{T} - \frac{\left(\sum_{k_0+m+1}^{T} y_{t-1}\varepsilon_t\right)^2}{T\sum_{k_0+m+1}^{T} y_{t-1}^2} \\
&= \frac{\sum_1^T \varepsilon_t^2}{T} + o_p(1) + o_p(1) + o_p(1) + \left(\hat{\beta}_1\left(\tau_0+\frac{m}{T}\right)-\beta_2\right)^2 \frac{\sum_{k_0+1}^{k_0+m} y_{t-1}^2}{T} \\
&\quad + o_p(1) \\
&\Rightarrow \sigma^2 + \frac{(1-\beta_2)\sigma^2 B^2(\tau_0)(1-\beta_2^{2m})}{1+\beta_2}.
\end{aligned}$$

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 Appendixes C and G. We want to prove a stronger result that

$\lim_{T\to\infty} \Pr(\hat{k} \neq k_0) = 0.$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

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,

$$\begin{aligned}\Pr(\hat{k} \neq k_0) &= \Pr(|\hat{k} - k_0| > M) + \Pr(|\hat{k} - k_0| \leq M \quad \text{and} \quad \hat{k} \neq k_0) \\ &\leq \eta + \Pr(|\hat{k} - k_0| \leq M \quad \text{and} \quad \hat{k} \neq k_0) \\ &\leq \eta + \sum_{m=1}^{M} \Pr\left(RSS_T\left(\tau_0 - \frac{m}{T}\right) - RSS_T(\tau_0) < 0\right) \\ &\quad + \sum_{m=1}^{M} \Pr\left(RSS_T\left(\tau_0 + \frac{m}{T}\right) - RSS_T(\tau_0) < 0\right).\end{aligned}$$

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 $\operatorname{plim} \hat{\beta}_1(\hat{\tau}_T) = \operatorname{plim} \hat{\beta}_1(\tau_0) = \beta_1 = 1$, $\operatorname{plim} \hat{\beta}_2(\hat{\tau}_T) = \operatorname{plim} \hat{\beta}_2(\tau_0) = \beta_2$. 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,

$$\begin{aligned}&\lim_{T\to\infty} \Pr(T(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \leq x) \\ &\quad = \lim_{T\to\infty} \Pr(T(\hat{\beta}_1(\tau_0) - \beta_1) \leq x) \lim_{T\to\infty} \Pr(\hat{k} = k_0) \\ &\qquad + \lim_{T\to\infty} \Pr(T(\hat{\beta}_1(\hat{\tau}_T) - \beta_1) \leq x | \hat{k} \neq k_0) \lim_{T\to\infty} \Pr(\hat{k} \neq k_0) \\ &\quad = \lim_{T\to\infty} \Pr(T(\hat{\beta}_1(\tau_0) - \beta_1) \leq x).\end{aligned}$$

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

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$T(\hat{\beta}_1(\hat{\tau}_T)-1) \overset{d}{=} T(\hat{\beta}_1(\tau_0)-1) = \frac{\sum_{1}^{[\tau_0 T]} \varepsilon_t y_{t-1}/T}{\sum_{1}^{[\tau_0 T]} y_{t-1}^2/T^2} \Rightarrow \frac{B^2(\tau_0)-\tau_0}{2\int_0^{\tau_0} B^2},$$

$$\sqrt{T}(\hat{\beta}_2(\hat{\tau}_T)-\beta_2) \overset{d}{=} \sqrt{T}(\hat{\beta}_2(\tau_0)-\beta_2) = \frac{\sum_{[\tau_0 T]+1}^{T} \varepsilon_t y_{t-1}/\sqrt{T}}{\sum_{[\tau_0 T]+1}^{T} y_{t-1}^2/T} \Rightarrow \frac{\sqrt{1-\beta_2^2}\bar{B}(1)}{1-\tau_0+B^2(\tau_0)},$$

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 β_1 at one and let $\beta_{2T} = 1 - [1/\sqrt{Tg(T)}]$, where $g(T) > 0$, with $g(T) \to \infty$ and $g(T)/\sqrt{T} \to 0$ as $T \to \infty$. Let υ be a finite constant and $B(\cdot)$, $B_1(\cdot)$, and $B_2(\cdot)$ be defined as in Section 2.

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

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

(ii) $$(\beta_{2T}-1)\frac{\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 \sum_{[\tau_0 T]+1}^{T} y_{t-1}\varepsilon_t}{\sum_{[\tau T]+1}^{T} y_{t-1}^2} = -\frac{O_p(Tg(T))O_p(T)}{\sqrt{Tg(T)}O_p(T^2)} = O_p\left(\sqrt{\frac{g(T)}{T}}\right)$$
$$= o_p(1);$$

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

(iv) $$\frac{1}{\sqrt{Tg(T)}} \sum_{0}^{[|\upsilon|g(T)]-1} y_{k_0-t-1}\varepsilon_{k_0-t}$$
$$= \frac{1}{\sqrt{Tg(T)}} \sum_{0}^{[|\upsilon|g(T)]-1} \left(y_{k_0} - \sum_{i=0}^{t} \varepsilon_{k_0-i}\right)\varepsilon_{k_0-t}$$
$$= \frac{y_{k_0}}{\sqrt{T}} \frac{1}{\sqrt{g(T)}} \sum_{0}^{[|\upsilon|g(T)]-1} \varepsilon_{k_0-t} - \frac{g(T)}{\sqrt{T}}$$
$$\times \sum_{0}^{[|\upsilon|g(T)]-1} \left(\frac{1}{\sqrt{g(T)}} \sum_{i=0}^{t} \varepsilon_{k_0-i}\varepsilon_{k_0-t}\right)\frac{1}{g(T)}$$
$$= -\frac{y_{k_0}}{\sqrt{T}} \frac{1}{\sqrt{g(T)}} \sum_{0}^{[|\upsilon|g(T)]-1} (-\varepsilon_{k_0-t}) - O_p\left(\frac{g(T)}{\sqrt{T}}\right) \Rightarrow -\sigma^2 B(\tau_0)B_1(-\upsilon);$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

$$\text{(v)}\quad \frac{1}{Tg(T)}\sum_{0}^{[|v|g(T)]-1} y_{k_0-t-1}^2$$
$$= \frac{1}{Tg(T)}\sum_{0}^{[|v|g(T)]-1}\left(y_{k_0}-\sum_{i=0}^{t}\varepsilon_{k_0-i}\right)^2$$
$$= |v|\frac{y_{k_0}^2}{T} - 2\sqrt{\frac{g(T)}{T}}\frac{y_{k_0}}{\sqrt{T}}\sum_{0}^{[|v|g(T)]-1}\left(\frac{1}{\sqrt{g(T)}}\sum_{i=0}^{t}\varepsilon_{k_0-i}\right)\frac{1}{g(T)}$$
$$+ \frac{g(T)}{T}\frac{1}{g(T)}\sum_{0}^{[|v|g(T)]-1}\left(\frac{1}{\sqrt{g(T)}}\sum_{i=0}^{t}\varepsilon_{k_0-i}\right)^2$$
$$= |v|\frac{y_{k_0}^2}{T} + O_p\left(\sqrt{\frac{g(T)}{T}}\right) + O_p\left(\frac{g(T)}{T}\right) \Rightarrow |v|\sigma^2 B^2(\tau_0).$$

Thus, equation (B.2) in Appendix B becomes

$$RSS_T(\tau) - RSS_T(\tau_0)$$
$$= -2(\beta_{2T}-1)\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}\varepsilon_t + (\beta_{2T}-1)^2\sum_{[\tau T]+1}^{[\tau_0 T]} y_{t-1}^2 + o_p(1)$$
$$= -\frac{2}{\sqrt{Tg(T)}}\sum_{0}^{[|v|g(T)]-1}(-y_{k_0-t-1}\varepsilon_{k_0-t}) + \frac{1}{g(T)T}\sum_{0}^{[|v|g(T)]-1} y_{k_0-t-1}^2 + o_p(1)$$
$$\Rightarrow -2\sigma^2 B(\tau_0)B_1(|v|) + |v|\sigma^2 B^2(\tau_0).$$

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

$$\text{(vi)}\quad \Lambda_T(\tau) = o_p(1),$$

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

$$\text{(viii)}\quad \frac{\sum_{1}^{[\tau_0 T]} y_{t-1}^2}{\sum_{1}^{[\tau T]} y_{t-1}^2} \xrightarrow{p} 1,$$

equation (B.4) in Appendix B becomes

$$RSS_T(\tau) - RSS_T(\tau_0)$$
$$= 2(\beta_{2T}-1)\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}\varepsilon_t + (\beta_{2T}-1)^2\sum_{[\tau_0 T]+1}^{[\tau T]} y_{t-1}^2 + o_p(1)$$
$$= -\frac{2}{\sqrt{Tg(T)}}\sum_{0}^{[vg(T)]-1} y_{k_0+t}\varepsilon_{k_0+t+1} + \frac{1}{Tg(T)}\sum_{0}^{[vg(T)]-1} y_{k_0+t}^2 + o_p(1)$$
$$= -2\frac{y_{k_0}}{\sqrt{T}}\frac{1}{\sqrt{g(T)}}\sum_{0}^{[vg(T)]-1}\varepsilon_{k_0+t+1} + v\frac{y_{k_0}^2}{T} + o_p(1)$$
$$\Rightarrow -2\sigma^2 B(\tau_0)B_2(v) + v\sigma^2 B^2(\tau_0).$$

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press

Applying the continuous mapping theorem for argmax functionals, we have

$$(1-\beta_{2T})^2T^2(\hat{\tau}_T-\tau_0)=\frac{T}{g(T)}(\hat{\tau}_T-\tau_0)=\hat{v}=\operatorname*{Arg\,min}_{v\in R}\{RSS_T(\tau)-RSS_T(\tau_0)\}$$

$$\xrightarrow{d}\operatorname*{Arg\,min}_{v\in R}\left\{-2\sigma^2B^2(\tau_0)\left(\frac{B^*(v)}{B(\tau_0)}-\frac{1}{2}|v|\right)\right\}$$

$$=\operatorname*{Arg\,max}_{v\in R}\left(\frac{B^*(v)}{B(\tau_0)}-\frac{1}{2}|v|\right),$$

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 $v>0$. ■

https://doi.org/10.1017/S0266466601171045 Published online by Cambridge University Press