Statistics and Its Interface Volume 9 (2016) 37–45

Sieve bootstrap monitoring persistence change in long memory process

Zhanshou Chen*, Zheng Tian, and Yuhong Xing

This paper adopts a moving ratio statistic to monitor persistence change in long memory process. The limiting distribution of monitoring statistic under the stationary long memory null hypothesis is derived. We show that the proposed monitoring scheme is consistent for stationary to nonstationary change. In particular, a sieve bootstrap approximation method is proposed. The sieve bootstrap method is used to determine the critical values for the null distribution of monitoring statistic which depends on unknown long memory parameter. The empirical size, power and average run length of the proposed monitoring procedure are evaluated in a simulation study. Simulations indicate that the new monitoring procedure performs well in finite samples. Finally, we illustrate our monitoring procedure using a set of foreign exchange rate data.

AMS 2000 subject classifications: Primary 60G22, 62F40; secondary 62L10.
Keywords and phrases: Change in persistence, Long memory process, Sieve bootstrap, Monitoring.

1. INTRODUCTION

Since the founding of regime switch between trend stationary and difference stationary by De Long and Summers (1988) [13], there is a growing body of evidence shown that economic and financial time series display changes in persistence. This has been an issue of substantial empirical interest, especially concerning inflation rate series, short-term interest rates, government budget deficits and real output. Recently, a number of testing procedures have been suggested that aim to distinguish such behavior. These include, *inter alia*, ratio tests (Kim (2000) [21]), LBI tests (Busetti and Taylor (2004) [1]), CUSUM of squares-based test (Leybourne et al. (2007a) [23]). More recently, Cavaliere and Taylor (2008) [3] considered persistence change test under non-stationary volatility innovation case, Sibbertsen and Kruse (2009) [28] studied long-range dependence innovation case, Hassler and Scheithauer (2011) [17] applied ratio tests and LBI tests to detect change point from short to long memory. Since multiple change-point test also is an important issue in change point analysis, Leybourne et al. (2007b) [24] proposed a multiple change points detection procedure based on sequences of doubly-recursive implementations of the regression-based unit root statistic of Elliott et al. (1996) [14]. Chen et al. (2012b) [7] proposed a moving ratio statistic to test multiple persistence changes in infinite variance linear processes. The finite variance case had been studied by Chen and Tian (2012a) [6] via a modified ratio statistic for the traditional ratio statistic is sensitive for variance change (Chen and Tian (2014) [9]). Kejriwal et al. (2012) [20] adopted a Wald statistic to test multiple structural changes in persistence. Martins and Rodrigues (2012) [25] and Hassler and Meller (2013) [18] considered multiple change point detecting problem in long memory process and illustrated their test procedure by some inflation rate series.

All of the above works are a retrospective test, namely, detect change point in a fixed historical sample. However, many economic and financial data arrive steadily and cheaply. It is desirable to sequentially detect whether the new arrival of data can be described by the current model or indicates that a change in the stochastic structure has taken place. This leads to a sequential test also which is an important issue in change point analysis. For surveys we refer the reader to Chu et al. (1996) [10], Gombay and Serban (2009) [16], Chen and Tian (2010b) [5] among many others. Steland (2007) proposed a kernel weighted variance ratio statistic to monitor change in persistence. Chen et al. (2012c) [8] extended the results to infinite variance case. For the kernel weighted variance ratio statistic does not provide a consistent monitoring procedure for $I(1)$ to $I(0)$ change, Chen et al. (2010a) [4] proposed a moving ratio statistic to overcome this problem. However, all these tests stay in the classical $I(1)/I(0)$ framework.

It is broadly accepted that many economic variables exhibit long-range dependencies that cannot be covered by the classical framework. Also in the more flexible I(d) framework, $0 \leq d < 3/2$, it is crucial to know whether the memory parameter is in the stationary or in the non-stationary region throughout or whether there is a change in persistence. Chen et al. (2012b) [7] showed by simulation that the moving ratio statistic also has power to distinguish such behavior, but it undergoes serious size distortions if one uses critical values which are obtained at the short memory case. In this paper, we derive the limiting distribution of our proposed moving ratio monitoring statistic under the stationary long memory null hypothesis. The derived limiting distri-

*Corresponding author.

butions are functionals of fractional Brownian motion depending on long memory parameter. We further prove the consistency of our monitoring procedure under the alternative hypothesis. In order to determine the critical values of the monitoring statistic which contains unknown long memory parameter, we propose a sieve bootstrap approximation method. The sieve bootstrap method was first introduced by Bühlmann (1997) [2]. Since then, many authors have studied this method. For surveys we refer the reader to Park (2002) [26] and references therein. Poskitt (2008) [27] studied the properties of the sieve bootstrap for fractionally integrated series and concluded that the sieve bootstrap is particularly useful in analyzing fractionally integrated process. Gerolimetto and Procidano (2008) [15] has studied fractional cointegration test using sieve bootstrap. Although Kirch (2008) [22] has argued that bootstrap method performs well to sequentially detect change point in online data, there are few papers that concentrate on sieve bootstrap for online long memory process. Our simulations indicate that the sieve bootstrap method proposed in this paper can not only circumvent the estimation of long memory parameter, but also can obtain asymptotically correct critical values of the moving ratio monitoring statistic. Finally, we illustrate our monitoring procedure via some Sweden/U.S. foreign exchange rate data.

The rest of the paper is organized as follows. Section 2 introduces the model and monitoring procedure. Section 3 shows the sieve bootstrap approximation method. In Section 4 we use Monte Carlo method to evaluate the finite sample performance of our monitoring procedure and sieve bootstrap method, an empirical application example also will be reported in this section. We conclude the paper in Section 5. All technical proofs of the theoretical results are gathered in the Appendix Section.

2. MONITORING PROCEDURE

Let $y_1, y_2, \ldots,$ be an observed time series that can be decomposed as

$$y_t = \mu_t + x_t, \quad (1-L)^d x_t = \varepsilon_t, \ t = 1, 2, \ldots, \tag{1}$$

where $\mu_t = E(y_t) = \delta' \gamma_t$ is a deterministic component modelled as a linear combination of a vector of nonrandom regressors γ_t. Typical components of γ_t are a constant, a time trend or dummy variables. This paper concentrates on $\gamma_t = 0$, $\gamma_t = 1$ and $\gamma_t = (1,t)^{\mathrm{T}}$ three different typical components. Namely, $\mu_t = 0$ if $\gamma_t = 0$, $\mu_t = \delta_0 \neq 0$ if $\gamma_t = 1$, and $\mu_t = \delta_0 + \delta_1 t$ with $\delta_1 \neq 0$ if $\gamma_t = (1,t)^{\mathrm{T}}$. δ_0 and δ_1 are unknown parameters. Random component x_t is the fractionally integrated process, in which L is the lag operator, ε_t are i.i.d. random variables with mean zero and variance σ^2. x_t has the MA representation

$$x_t = \sum_{j=0}^{\infty} w_j(d) \varepsilon_{t-j} \tag{2}$$

where, letting $\Gamma(\cdot)$ denotes the gamma function,

$$w_j(d) = \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(1+j)}.$$

The degree of integration of y_t is therefore solely determined by the memory parameter d. The memory parameter is restricted to $0 \leq d < 3/2$. Note that the process y_t is a stationary long memory process if $0 < d < 1/2$, the process y_t is a non-stationary long memory process if $1/2 < d < 3/2$. To simplify the notation we denote $y_t \sim I(d_1)$ with $0 < d_1 < 1/2$ if y_t is a stationary long memory process, and $y_t \sim I(d_2)$ with $1/2 < d_2 < 3/2$ if y_t is a non-stationary long memory process.

We focus on the following change point problem: observe sequences $y_1, y_2, \ldots,$ and detect whether a stationary to non-stationary change occurs in long memory process (1), namely, we want to test the null hypothesis

$$H_0 : y_t \sim I(d_1), \ t = 1, 2, \ldots, T, \tag{3}$$

against the alternative hypothesis

$$\begin{aligned} H_1 : \quad & y_t \sim I(d_1), \ t = 1, \ldots, k^*, \\ & y_t \sim I(d_2), \ t = k^* + 1, \ldots, T. \end{aligned} \tag{4}$$

where T denotes the largest monitoring sample size, k^* is the unknown change point. Our monitoring procedure is based on the following assumption:

Assumption 2.1. Assume $y_t \sim I(d_1), \ t = 1, \ldots, [T\tau]$, $\tau \in (0,1)$.

This assumption is similar to the "noncontamination assumption" of Chen and Tian (2010b) [5] who monitors parameter changes in linear regression models. We start from the $([T\tau]+1)st$ sample sequentially detect $I(d_1)$ to $I(d_2)$ change until time horizon T by the following moving ratio statistic

$$\Gamma_T(s) = \frac{\sum_{t=[Ts]-[T\tau]+1}^{[Ts]} (\sum_{i=[Ts]-[T\tau]+1}^{t} \hat{\varepsilon}_{1,i})^2}{\sum_{t=1}^{[T\tau]} (\sum_{i=1}^{t} \hat{\varepsilon}_{0,i})^2}, \tag{5}$$

where $[x]$ denotes the largest integer smaller than x, $\hat{\varepsilon}_{1,i}$ represents the OLS residuals from the regression of y_t on γ_t, $t = [Ts] - [T\tau] + 1, \ldots, [Ts]$, and $\hat{\varepsilon}_{0,i}$ represents the OLS residuals from the regression of y_t on γ_t, $t = 1, \ldots, [T\tau]$.

The denominator and numerator of statistic (5) have the same converge rate if y_t keep $I(d_1)$ process. On the contrary, the denominator has a faster converge rate than the numerator if y_t occur a change from $I(d_1)$ to $I(d_2)$ at time $t > [T\tau]$. This causes the statistic (5) to diverge to infinity. So, we can define the stopping time as

$$R_T(n) = \min\{[T\tau] < n \leq T : \Gamma_T(n/T) > c\}, \tag{6}$$

with the convention $\min\{\Phi\} = T$ for some critical value c.

Before deriving the asymptotic distribution of statistic (5), we have to define type I fractional Brownian motion (see Davidson and Hashimzade (2009) [12])

$$W_d(r) = \frac{1}{\Gamma(d+1)}\int_0^r (r-s)^d dB(s) + \frac{1}{\Gamma(d+1)}\int_{-\infty}^0 [(r-s)^d - (-s)^d] dB(s),$$

where $-\frac{1}{2} < d < \frac{1}{2}$ and B denotes regular Brownian motion.

Theorem 2.1. *If Assumption 2.1 holds, then under the null hypothesis (3)*

$$\Gamma_T(s) = \Upsilon(s) \Rightarrow \frac{\int_{s-\tau}^s (U_{j,0}(d_1,u))^2 du}{\int_0^\tau (U_{j,1}(d_1,u))^2 du},$$

where $\Rightarrow$ denotes weak convergence, $U_{j,0}(d_1,u)$ and $U_{j,1}(d_1,u)$, $j = 1,2,3$ are functionals of type I fractional Brownian motion $W_{d_1}(\cdot)$, $j = 1$ denotes $\gamma_t = 0$, $j = 2$ denotes $\gamma_t = 1$ and $j = 3$ denotes $\gamma_t = (1,t)^{\mathrm{T}}$,

$$\begin{aligned}
U_{1,0}(d_1,u) &= W_{d_1}(u);\\
U_{1,1}(d_1,u) &= W_{d_1}(u) - W_{d_1}(s-\tau);\\
U_{2,0}(d_1,u) &= W_{d_1}(u) - u\tau^{-1}W_{d_1}(\tau);\\
U_{2,1}(d_1,u) &= \frac{s-u}{\tau}(W_{d_1}(s) - W_{d_1}(s-\tau))\\
&\quad - (W_{d_1}(s) - W_{d_1}(u))\,;\\
U_{3,0}(d_1,u) &= W_{d_1}(u) - \frac{3u^2(2-\tau) - 2u\tau(3-\tau)}{2\tau^3}W_{d_1}(\tau)\\
&\quad - \frac{3u(u-\tau)}{\tau^3}\int_0^\tau W_{d_1}(v)dv;\\
U_{3,1}(d_1,u) &= W_{d_1}(u-s+\tau)\\
&\quad - 6\tau(s-u)K\int_{s-\tau}^s W_{d_1}(v)dv\\
&\quad - \tau[6s(s-u-1) + 6u]KW_{d_1}(\tau)\\
&\quad + \tau^2(\tau - 3s + 3u)KW_{d_1}(\tau);
\end{aligned}$$

and

$$K = \frac{u-s+\tau}{4\tau\left(s^3 - (s-\tau)^3\right) - 3\tau^2(2s-\tau)^2}. \tag{7}$$

Theorem 2.1 implies that

$$R_T(s) \xrightarrow{d} \min\{\tau < s \le 1 : \Upsilon(s) > c\}, \quad \text{as } T \to \infty.$$

Theorem 2.1 supplies the asymptotic distribution of monitoring statistic $\Gamma_T(s)$ when innovation process $\{\varepsilon_t, -\infty < t < \infty\}$ in model (1) are i.i.d. In fact, this result is still hold when innovation process $\{\varepsilon_t\}$ satisfies the following Assumption 2.2. For more detail discussions about this assumption, we refer the reader to Davidson and De Jong (2000) [11].

Assumption 2.2. The mean zero sequence $\{\varepsilon_t\}$

(1) is uniformly L_r-bounded for $r > 2$,
(2) is L_2-NED of size $-1/2$ on V_t with $d_t = 1$, where V_t is either an α-mixing sequence of size $-r/(r-2)$, or a ϕ-mixing sequence of size $-r/(2(r-1))$,
(3) is covariance stationary, and

$$0 < \sigma_\varepsilon^2 = \lim_{n\to\infty} n^{-1}\sum_{t=1}^n\sum_{s=1}^n E(\varepsilon_t\varepsilon_s) < \infty.$$

Corollary. *If the innovation process $\{\varepsilon_t, -\infty < t < \infty\}$ in model (1) satisfies Assumption 2.2, then Theorem 2.1 still holds.*

The following Theorem 2.2 shows the consistency of moving ratio statistic (5) for stationary to non-stationary persistence change.

Theorem 2.2. *Under Assumption 2.1 or 2.2, if there occurs a change from $I(d_1)$ to $I(d_2)$ at $[T\tau^*]$, then we have*

$$\Gamma_T(s) = O_p(T^{2(d_2-d_1)}), s \in (\tau^*, 1].$$

3. SIEVE BOOTSTRAP APPROXIMATION

One drawback of statistic $\Gamma_T(s)$ is that having asymptotic distribution depends on long memory parameter d_1. In order to avoid estimating this nuisance parameter, we now resort to the bootstrap methodology. The object of this section is to develop an approximation to the null distribution of statistic $\Gamma_T(s)$, even if long memory parameter is unknown. We will now explain how the sieve bootstrap is applied in our context.

The steps of our sieve bootstrap methodology are constructed as follows:

Step 1. Having observed the first $[T\tau]$ samples of training sample $y_1, y_2, \ldots, y_{[T\tau]}$, compute the OLS residuals $\hat{x}_t = y_t - \hat{\delta}'\gamma_t$.

Step 2. We fit autoregressive processes to the residuals

$$\hat{x}_t = \beta_1\hat{x}_{t-1} + \beta_2\hat{x}_{t-2} + \cdots + \beta_{p(T)}\hat{x}_{t-p(T)} + \varepsilon_t$$

with fixed $p(T) = 10\log_{10}([T\tau])$ and white nose ε_t, then, choose an optimal p using the AIC criterion, $p = p_{AIC}$. Then, we obtain the estimated coefficients $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_{[T\tau]}$ via the Yule-Walker equation.

Step 3. Using $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p$ we construct a set of residuals from the fitted process

$$\hat{\varepsilon}_t = \sum_{j=0}^p \hat{\beta}_j\hat{x}_{t-j}, \quad \hat{\beta}_0 = 1,\ p+1 \le t \le [T\tau].$$

Step 4. Compute the centered residuals

$$\tilde{\varepsilon}_t = \hat{\varepsilon}_t - \frac{1}{T-p}\sum_{i=p+1}^{[T\tau]}\hat{\varepsilon}_i$$

Step 5. Select with replacement a sieve bootstrap residuals $\{\varepsilon_j^*, j = 1, \ldots, T\}$ from $\{\tilde{\varepsilon}_j, j = p+1, \ldots, [T\tau]\}$.

Step 6. Generate sieve bootstrap observations x_t^* using ε_j^* according to

$$x_t^* = \hat{\beta}_1 x_{t-1}^* + \hat{\beta}_2 x_{t-2}^* + \cdots + \hat{\beta}_p x_{t-p}^* + \varepsilon_t^*, \quad t = 1, \ldots, T,$$
$$y_t^* = \hat{\delta}' \gamma_t + x_t^*,$$

where the order p was chosen in Step 2, and the starting p observations $x_{p-1}^*, x_{p-2}^*, \ldots, x_0^*$ can be set equal to $\bar{y}_{[T\tau]} = \frac{1}{[T\tau]} \sum_{t=1}^{[T\tau]} y_t$.

Step 7. Calculate the statistic

$$\Gamma_T^*(s) = \frac{\sum_{t=[Ts]-[T\tau]+1}^{[Ts]} (\sum_{i=[Ts]-[T\tau]+1}^{t} \hat{\varepsilon}_{1,i}^*)^2}{\sum_{t=1}^{[T\tau]} (\sum_{i=1}^{t} \hat{\varepsilon}_{0,i}^*)^2},$$

where $\hat{\varepsilon}_{1,i}^*$ denotes the OLS residuals from the regression of y_t^* on γ_t, $t = [Ts] - [T\tau] + 1, \ldots, [Ts]$, $\hat{\varepsilon}_{0,i}^*$ denotes the OLS residuals from the regression of y_t^* on γ_t, $t = 1, \ldots, [T\tau]$.

Step 8. Repeat Step 5 to Step 7 B times, approximate the asymptotic critical value of statistic $\Gamma_T(s)$ by the empirical quantile of $\Gamma_T^*(s)$.

Remark: Kapetanios and Psaradakis (2006) [19] have established this procedure's asymptotic validity in approximating sample mean and covariance of long memory process. The advantage of the above sieve bootstrap is that we only have to calculate the critical values once at the beginning of the monitoring period. Although we have only few historical data to base on a bootstrap for a much longer time series, simulations which will be listed in the next section show that this sieve bootstrap controls empirical size well and gives satisfactory empirical power. We do not construct sieve bootstrap procedure based on all data available up to monitoring time point or updated bootstrap procedure which has been proposed by Steland (2006) [29] not only because these procedures are computationally very expensive, but also they will lose some test power.

4. SIMULATIONS AND EMPIRICAL APPLICATION

4.1 Simulations

In this section, we use Monte Carlo methods to investigate the finite sample performance of our monitoring procedure introduced in Section 2 and sieve bootstrap approach proposed in Section 3. First, we consider the following data generating process (DGP)

$$(8) \quad y_t = r_0 + r_1 t + x_t, \quad x_t = \sum_{j=0}^{J} w_{t-j}(d) \varepsilon_j, \; t = 1, \ldots, T,$$

where r_0, d are parameters, and J is some given constant. $\{\varepsilon_t\}$ are *i.i.d.* $N(0,1)$ innovations and

$$(9) \qquad w_0(d) = 1, \quad w_j(d) = \frac{j + d - 1}{j} \cdot w_{j-1}(d).$$

The values of parameters in model (8) were chosen to be $J = 1{,}000$, long memory parameter d varying among 0.1, 0.3, $r_0 = 0.1$ for de-meaned data, and $r_0 = r_1 = 0.1$ for de-trended data. Throughout this section we fix the sieve bootstrap frequency $B = 500$. The start time was set to be $\tau = 0.2$, 0.3, sample size $T = 200$, 500, and all simulations are obtained by 1,000 replications at $\alpha = 5\%$ nominal level.

Table 1 reports the empirical sizes of monitoring statistic $\Gamma_T(s)$ both asymptotic and bootstrap for the de-meaned data. The empirical sizes for the de-trended data are listed in Table 2. The asymptotic critical values are obtained via direct simulation from 1,0000 repetitions by setting T = 1,000 in model (8). From these two tables we can see that both tests exhibit similar size distortions, and all these size distortions becoming light as sample size or start time increases. Although the value of long memory parameter also has some influence to the empirical size, we can not conclude that a larger value gives more severe size distortion. This indicates that our proposed sieve bootstrap procedure in Section 3 gives asymptotic correct critical values for statistic $\Gamma_T(s)$ under the stationary long memory null hypothesis. This supports the motivation of Section 3. Since the size distortion of statistic $\Gamma_T(s)$ is still not significant if $d = 0$, we say that our proposed sieve bootstrap method is still robust for short memory process. Compare Table 1 and Table 2 we can see, it gives lighter size distortion for the de-meaned data. We think it mainly because one has to estimate one more unknown parameter under the de-trended case.

To investigate the alternative performance of the proposed monitoring procedure and the sieve bootstrap approximation method, we consider the following DGP

$$(10) \quad \begin{aligned} y_t &= r_0 + r_1 t + x_t, \\ x_t &= \begin{cases} x_{k^*+1} + \sum_{j=0}^{J} w_j(d_1) e_{t-j}, \; t = 1, 2, \ldots, k^*, \\ \sum_{j=0}^{J} w_j(d_2) e_{t-j}, \; t = k^* + 1, \; k^* + 2, \ldots, T, \end{cases} \end{aligned}$$

where $r_0, r_1, w_j(d)$ and e_t have same definitions as in (9). DGP (10) shows that the long memory parameter d_1 changes to d_2 at change point k^*. We add the value x_{k^*+1} in the first k^* samples. We just want to delete the sharp change point at k^*. To evaluate the influence of change point location, we are varying k^* among $0.3T, 0.5T$ and $0.7T$. The empirical powers and average run length (ARL) which is defined as the average sample number between change time and giving signal time, for the de-meaned and de-trended data are presented in Table 3 and Table 4 respectively. From these tables, we can see that the empirical power increases as sample size or start time increases. This is an expected result for our monitoring procedure which is a consistent test. Further, a smaller start time and larger change size provide shorter ARL, larger distance between start time and location of change often causing the signal to come late. These conclusions are similar to Chen et al. (2010) [4] who obtained it under the short memory null hypothesis. Although

Table 1. Empirical sizes with de-meaned data

		$\tau = 0.2$			$\tau = 0.3$		
	T	$d = 0$	$d = 0.1$	$d = 0.3$	$d = 0$	$d = 0.1$	$d = 0.3$
Bootstrap	200	0.047	0.054	0.057	0.054	0.061	0.045
	500	0.050	0.048	0.052	0.048	0.059	0.049
Asymptotic	200	0.053	0.055	0.058	0.052	0.044	0.055
	500	0.049	0.047	0.051	0.050	0.052	0.046

Table 2. Empirical sizes with de-trended data

		$\tau = 0.2$			$\tau = 0.3$		
	T	$d = 0$	$d = 0.1$	$d = 0.3$	$d = 0$	$d = 0.1$	$d = 0.3$
Bootstrap	200	0.054	0.057	0.060	0.051	0.060	0.055
	500	0.055	0.053	0.058	0.052	0.062	0.064
Asymptotic	200	0.047	0.054	0.058	0.050	0.043	0.056
	500	0.051	0.052	0.053	0.048	0.051	0.058

Table 3. Empirical powers and ARL with de-meaned data

		$k^* = 0.3$		$k^* = 0.5$		$k^* = 0.7$	
T	d	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.2$	$\tau = 0.3$
200	$0.1 \to 0.6$	0.318	0.409	0.260	0.328	0.186	0.246
		[57.2]	[65.3]	[39.6]	[47.6]	[14.7]	[21.7]
	$0.1 \to 0.9$	0.629	0.765	0.576	0.681	0.441	0.499
		[49.9]	[51.5]	[37.0]	[43.6]	[22.4]	[27.0]
	$0.3 \to 0.6$	0.188	0.229	0.163	0.204	0.125	0.153
		[57.2]	[67.1]	[34.7]	[40.5]	[3.26]	[16.8]
	$0.3 \to 0.9$	0.447	0.573	0.388	0.495	0.294	0.337
		[54.2]	[59.8]	[38.4]	[44.7]	[17.0]	[27.2]
500	$0.1 \to 0.6$	0.491	0.563	0.423	0.503	0.306	0.375
		[132]	[144]	[101]	[105]	[52.6]	[67.7]
	$0.1 \to 0.9$	0.879	0.951	0.851	0.884	0.708	0.791
		[96.9]	[94.2]	[82.2]	[84.8]	[55.0]	[65.9]
	$0.3 \to 0.6$	0.255	0.327	0.193	0.253	0.153	0.201
		[135]	[156]	[103]	[110]	[38.8]	[51.4]
	$0.3 \to 0.9$	0.724	0.820	0.603	0.739	0.509	0.518
		[116]	[129]	[93.2]	[106]	[56.5]	[66.6]

it gives lower empirical power and longer ARL for the detrended data than the de-meaned data, the main conclusions are similar. This result is similar to the results of Sibbertsen and Kruse (2009) [28] who obtained it in the retrospective test.

In order to check whether our proposed monitoring procedure still works in a more general model, we also perform some simulation by set x_t in model (8) follow

$$(1 - 0.5L)(1 - B)^d x_t = \varepsilon_t, \tag{11}$$

and

$$(1 - B)^d x_t = (1 + 0.5L)\varepsilon_t. \tag{12}$$

DGP (11) is an ARFIMA(1,d,0) model, and DGP (12) is an ARFIMA(0,d,1) model. The other parameters have the same assumptions as in model (8) and (10). The empirical size and power, ARL at sample size $T = 200$ for de-meaned data are reported in Table 5 and Table 6 respectively. From Table 5 we can see that our monitoring procedure can still control the empirical size well for a small long memory parameter value. The size distortion becomes sever for a larger value of long memory parameter. Unreported simulations indicate that a larger sample size has little influence to control the empirical size well. Although we can conclude, by comparing the results in Table 6 with Table 3, that our monitoring procedure gives a better alternative performances when DGP follows an ARFIMA(0,d,0) model than other ARFIMA(p,d,q) model, the main conclusions are not changed.

In conclusion, our proposed moving ratio statistic and sieve bootstrap approximation method supply an efficient monitoring procedure to sequentially detect stationary to non-stationary regime switch for ARFIMA(0,d,0) process. Many times, the change point can be found very early and it is not necessary to wait until the time horizon. Unfor-

Table 4. Empirical powers and ARL with de-trended data

T	d	$k^*=0.3$ $\tau=0.2$	$k^*=0.3$ $\tau=0.3$	$k^*=0.5$ $\tau=0.2$	$k^*=0.5$ $\tau=0.3$	$k^*=0.7$ $\tau=0.2$	$k^*=0.7$ $\tau=0.3$
200	$0.1 \to 0.6$	0.226	0.319	0.198	0.2478	0.277	0.302
		[50.3]	[61.2]	[34.6]	[45.2]	[13.8]	[18.7]
	$0.1 \to 0.9$	0.519	0.632	0.474	0.576	0.542	0.601
		[45.5]	[52.1]	[33.3]	[41.0]	[20.8]	[24.8]
	$0.3 \to 0.6$	0.159	0.189	0.151	0.164	0.235	0.273
		[50.2]	[61.2]	[24.4]	[31.7]	[25.6]	[24.8]
	$0.3 \to 0.9$	0.387	0.503	0.298	0.425	0.354	0.411
		[47.4]	[50.9]	[34.4]	[39.8]	[26.6]	[35.0]
500	$0.1 \to 0.6$	0.433	0.502	0.383	0.447	0.288	0.351
		[135]	[150]	[97.6]	[104]	[46.6]	[71.7]
	$0.1 \to 0.9$	0.769	0.883	0.802	0.878	0.854	0.921
		[87.2]	[95.5]	[85.2]	[89.6]	[48.0]	[58.7]
	$0.3 \to 0.6$	0.215	0.286	0.188	0.244	0.250	0.310
		[128]	[158]	[100]	[112]	[40.6]	[54.4]
	$0.3 \to 0.9$	0.608	0.736	0.570	0.704	0.649	0.758
		[119]	[131]	[88.2]	[101]	[61.5]	[70.2]

Table 5. Empirical sizes under AEFIMA(0,d,1) and AEFIMA(1,d,0) model

Model	$\tau=0.2$ $d=0$	$\tau=0.2$ $d=0.1$	$\tau=0.2$ $d=0.3$	$\tau=0.3$ $d=0$	$\tau=0.3$ $d=0.1$	$\tau=0.3$ $d=0.3$
AEFIMA(0,d,1)	0.050	0.056	0.083	0.051	0.056	0.087
AEFIMA(1,d,0)	0.047	0.060	0.078	0.046	0.059	0.081

Table 6. Empirical powers and ARL under AEFIMA(0,d,1) and AEFIMA(1,d,0) model

Model	d	$k^*=0.3$ $\tau=0.2$	$k^*=0.3$ $\tau=0.3$	$k^*=0.5$ $\tau=0.2$	$k^*=0.5$ $\tau=0.3$	$k^*=0.7$ $\tau=0.2$	$k^*=0.7$ $\tau=0.3$
AEFIMA(0,d,1)	$0.1 \to 0.6$	0.304	0.380	0.277	0.347	0.200	0.242
		[55.8]	[64.9]	[39.7]	[47.1]	[19.9]	[22.6]
	$0.1 \to 0.9$	0.640	0.736	0.533	0.688	0.412	0.511
		[51.2]	[52.8]	[38.0]	[41.5]	[21.2]	[29.2]
	$0.3 \to 0.6$	0.212	0.247	0.210	0.209	0.131	0.168
		[57.4]	[62.2]	[33.4]	[42.6]	[5.61]	[13.8]
	$0.3 \to 0.9$	0.490	0.578	0.427	0.534	0.333	0.383
		[53.6]	[57.7]	[39.3]	[43.5]	[19.5]	[25.1]
AEFIMA(1,d,0)	$0.1 \to 0.6$	0.289	0.391	0.259	0.354	0.203	0.255
		[54.3]	[65.7]	[40.0]	[46.6]	[20.1]	24.4]
	$0.1 \to 0.9$	0.626	0.741	0.509	0.702	0.399	0.535
		[49.7]	[60.0]	[34.5]	[42.3]	[22.2]	[30.8]
	$0.3 \to 0.6$	0.194	0.251	0.187	0.213	0.155	0.183
		[55.1]	[63.3]	[32.8]	[43.0]	[8.43]	[15.5]
	$0.3 \to 0.9$	0.500	0.596	0.413	0.552	0.329	0.404
		[50.9]	[58.4]	[38.9]	[44.5]	[18.6]	[27.2]

tunately, our monitoring procedure still works when DGP follows a more general ARFIMA(p,d,q) model, but the size distortions are not neglectable for a large value of long memory parameter. In this case, a more robust monitoring procedure has to be designed. Recall that the monitoring procedure is more robust and powerful under a larger start time, whereas a smaller start time gives shorter ARL. An ideal way is that we can choose a larger start time for smaller monitoring horizon T, and choose a smaller start time for larger monitoring horizon. Although we can not find an optimum choice about start time, simulations indicate that the start time $\tau = 0.3$ is an acceptable choice to obtain satisfactory power and ARL if the monitoring horizon T is not too large or small. Of course, because the starting time has little influence to the empirical size, other small starting times also are acceptable if the sample size is large enough.

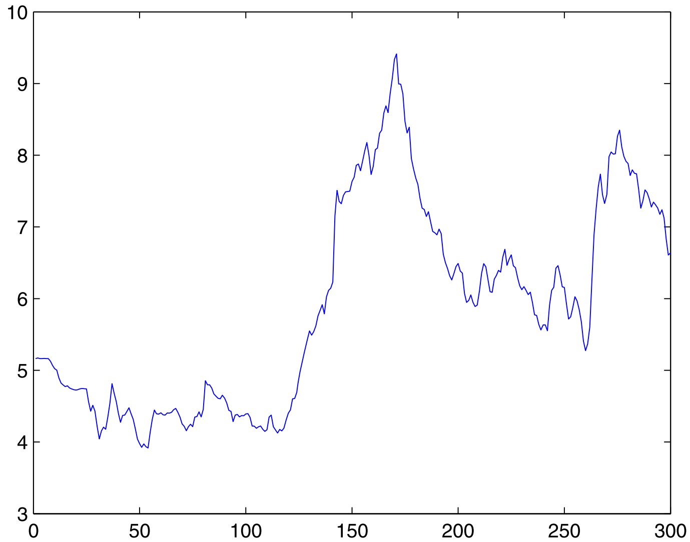

Figure 1. Sweden/U.S. foreign exchange rate monthly data from January 1, 1971 to December 1, 1995.

4.2 Empirical applications

In this section, we illustrate our monitoring procedure using a set of Sweden/U.S. foreign exchange rate monthly data which was observed from January 1, 1971 to December 1, 1995 with samples of 300 observations. This data set was download on the web site of the Federal Reserve Bank of St. Louis. Many papers have studied such a foreign exchange rate data under short memory or long memory condition. Fig. 1 reports the original data set.

We apply the monitoring statistic $\Gamma_T(s)$ and sieve bootstrap method monitor stationary to non-stationary persistence change by setting $B = 500$, $\tau = 0.3$, and nominal level $\alpha = 5\%$. The monitoring procedure stopped at observation 156 for de-meaned data and 158 for de-trend data. These results indicate that there occurs a stationary to non-stationary persistence change before October, 1961. Chen et al. (2012c) [8] has studied this data set using a kernel weighted variance ratio statistic and found that there is a stationary to non-stationary persistence change at observation 104. This coincides with our monitoring result.

5. CONCLUSIONS

This paper concentrates on persistence change point monitoring problem in long memory process. We adopt a moving ratio statistic sequentially to detect stationary to non-stationary persistence change. The null distribution of monitoring statistic and its consistency are proved. In order to avoid estimating long memory parameter in null distribution, we propose a sieve bootstrap approximation method. Empirical sizes in simulation study indicate that the sieve bootstrap method can supply asymptotic correct critical values. We also evaluated the empirical power and average run length of proposed monitoring procedures and found that our new method supplies an efficient monitoring procedure to sequentially detect stationary to non-stationary regime switch for long memory process. At last, we illustrate our monitoring procedure using a set of Sweden/U.S. foreign exchange rate monthly data.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No. 11301291, 61463045), Natural Science Foundation of Qinghai Province (No. 2012-Z-929Q) and the Science and Technology Innovation Foundation of Qinghai Normal University (2012).

APPENDIX SECTION

Proof of Theorem 2.1. Recall that if time series x_t follows a stationary long memory process with i.i.d. innovation and long memory parameter $0 < d_1 < 1/2$, then

$$T^{-1/2-d_1}\sum_{i=1}^{[T\tau]} x_i \Rightarrow \omega W_{d_1}(r), \tag{13}$$

where $W_{d_1}(\cdot)$ is the type I fractional Brownian motion with self-similarity parameter d_1, ω^2 is the long-run variance. Let $t = [Tu]$, $i = [Tv]$, then if $\gamma_t = 0$,

$$\begin{aligned} T^{-1/2-d_1}\sum_{i=1}^{t}\hat{\varepsilon}_{0,i} &= T^{-1/2-d_1}\sum_{i=1}^{t} x_i \\ &\Rightarrow \omega W_{d_1}(v) \equiv \omega U_{1,0}(d_1,u), \end{aligned} \tag{14}$$

$$\begin{aligned} T^{-1/2-d_1}\sum_{i=[Ts]-[T\tau]+1}^{t}\hat{\varepsilon}_{1,i} &= T^{-1/2-d_1}\sum_{i=[Ts]-[T\tau]+1}^{t} x_i \\ &\Rightarrow \omega(W_{d_1}(v) - W_{d_1}(s-u)) \\ &\equiv \omega U_{1,1}(d_1,u). \end{aligned} \tag{15}$$

If $\gamma_t = 1$, observing that $\hat{\varepsilon}_{0,i} = x_i - [T\tau]^{-1}\sum_{j=1}^{[T\tau]} x_j$, $\hat{\varepsilon}_{1,i} = x_i - [T\tau]^{-1}\sum_{j=[Ts]-[T\tau]+1}^{[Ts]} x_j$. Thus,

$$\begin{aligned} & T^{-1/2-d_1}\sum_{i=1}^{t}\hat{\varepsilon}_{0,i} \\ &= T^{-1/2-d_1}\sum_{i=1}^{t} x_i - \frac{tT^{-1/2-d_1}}{[T\tau]}\sum_{j=1}^{[T\tau]} x_j \\ &\Rightarrow \omega(W_{d_1}(u) - u\tau^{-1}W_{d_1}(\tau)). \\ &\equiv \omega U_{2,0}(d_1,u). \end{aligned} \tag{16}$$

$$\begin{aligned} & T^{-1/2-d_1}\sum_{i=[Ts]-[T\tau]+1}^{t}\hat{\varepsilon}_{1,i} \\ &= T^{-1/2-d_1}\sum_{i=[Ts]-[T\tau]+1}^{t} x_i \end{aligned} \tag{17}$$

$$-\frac{t-[Ts]+[T\tau]}{[T\tau]T^{1/2+d_1}}\sum_{j=[Ts]-[T\tau]+1}^{[Ts]}x_j$$
$$=-T^{-1/2-d_1}\sum_{i=[Tu]+1}^{[Ts]}x_i$$
$$+\frac{[Ts]-[Tu]}{[T\tau]T^{1/2+d_1}}\sum_{j=[Ts]-[T\tau]+1}^{[Ts]}x_j$$
$$\Rightarrow \omega\left(\frac{s-u}{\tau}(W_{d_1}(s)-W_{d_1}(s-\tau))\right)$$
$$-\omega\left(W_{d_1}(s)-W_{d_1}(u)\right)$$
$$\equiv \omega U_{2,1}(d_1,u)$$

If $\gamma_t=(1,t)^{\mathrm{T}}$, let $\delta=(\alpha,\beta)^{\mathrm{T}}$, then by the definition of LS we have

$$\begin{pmatrix}\hat{\alpha}-\alpha\\ \hat{\beta}-\beta\end{pmatrix}=\begin{pmatrix}\sum & \sum t\\ \sum t & \sum t^2\end{pmatrix}^{-1}\begin{pmatrix}\sum x_t\\ \sum tx_t\end{pmatrix},$$

where $\sum=\sum_{t=1}^{[T\tau]}$, if we estimate δ using the samples $y_1,\ldots,y_{[T\tau]}$, and $\sum=\sum_{t=[Ts]-[T\tau]+1}^{[Ts]}$, if we estimate δ using the samples $y_{[Ts]-[T\tau]+1},\ldots,y_{[Ts]}$. Hence, by a tedious calculation we have

$$\text{(18)}\quad T^{-1/2-d_1}\sum_{i=1}^{t}\hat{\varepsilon}_{0,i}$$
$$=T^{-1/2-d_1}\sum_{i=1}^{[Tu]}\left(x_i-(\hat{\alpha}-\alpha)-(\hat{\beta}-\beta)i\right)$$
$$\Rightarrow \omega(W_{d_1}(u)-\frac{3u^2(2-\tau)-2u\tau(3-\tau)}{2\tau^3}W_{d_1}(\tau)$$
$$-\frac{3u(u-\tau)}{\tau^3}\int_0^{\tau}W_{d_1}(v)dv)$$
$$\equiv \omega^2V_{3,0}(d_1,u).$$

Let

$$K=\frac{u-s+\tau}{4\tau\left(s^3-(s-\tau)^3\right)-3\tau^2(2s-\tau)^2},$$

then, a similar arguments gives that

(19)

$$T^{-1/2-d_1}\sum_{i=1}^{t}\hat{\varepsilon}_{1,i}$$
$$\Rightarrow \omega(W_{d_1}(u-s+\tau)-6\tau(s-u)K\int_{s-\tau}^{s}W_{d_1}(v)dv$$
$$-\tau[6s(s-u-1)+\tau(\tau-3s+3u)+6u]KW_{d_1}(\tau))$$
$$\equiv \omega V_{3,1}(d_1,u),$$

Combining (14)–(19), using the continuous mapping theorem and the continuity of functionals of $U_{j,0}(d_1,u)$ and $U_{j,1}(d_1,u)$, $j=1,2,3$, we complete the proof of Theorem 2.1.

Proof of Corollary. Davidson and De Jong (2000) [11] showed that if innovation process satisfies Assumption 2.2, then the limiting distribution in (13) still holds. So, we can prove this Corollary using the same proof line as Theorem 2.1. We omit to report here again.

For the remainder of this section we omit proofs for the $\gamma_t=0$ and $\gamma_t=(1,t)^{\mathrm{T}}$ cases; these are straightforward but tedious and follow the same logical development as those presented for the $\gamma_t=1$ case.

Proof of Theorem 2.2. From the proof of Theorem 2.1 we have the denominator of $\Gamma_T(s)$ is $O_p(T^{2+2d_1})$. If $s>\tau^*$, then

$$T^{-1/2-d_2}\sum_{i=[Ts]-[T\tau]+1}^{[Ts]}x_i$$
$$=T^{-1/2-d_2}\sum_{i=[Ts]-[T\tau]+1}^{[T\tau^*]}x_i+T^{-1/2-d_2}\sum_{i=[T\tau^*]+1}^{[Ts]}x_i$$
$$\Rightarrow O_p(T^{d_1-d_2})+\omega\int_{\tau^*}^{s}W_{d_2}(u)du.$$

This indicates the numerator of $\Gamma_T(s)$ is $O_p(T^{2+2d_2})$ if $s>\tau^*$. Hence

$$\Gamma_T(s)=O_p(T^{2(d_2-d_1)}),\ s\in(\tau^*,1].$$

Received 10 March 2014

REFERENCES

[1] Busetti, F. and Taylor, A. M. R. (2004). Tests of stationarity against a change in persistence. *Journal of Econometrics*, 123, 33–66. MR2125437
[2] Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli*, 3(2), 123–148. MR1466304
[3] Cavaliere, G. and Taylor, A. M. R. (2008). Testing for a change in persistence in the presence of a non-stationary volatility. *Journal of Econometrics*, 147, 84–98. MR2472983
[4] Chen, Z., Tian, Z. and Wei, Y. (2010). Monitoring change in persistence in linear time series. *Statistics & Probability Letters*, 80, 1520–1527. MR2669755
[5] Chen, Z. and Tian, Z. (2010b). Modified procedures for change point monitoring in linear regression models. *Mathematics and Computers in Simulation*, 81, 62–75. MR2729128
[6] Chen, Z. and Tian, Z. (2012a). Moving ratio test for multiple changes in persistence, *Journal of Systems Science & Complexity*, 25(3), 582–593. MR2930561
[7] Chen, Z., Jin, Z., Tian, Z. and Qi, P. (2012b). Bootstrap testing multiple changes in persistence for a Heavy-tailed sequence. *Computational Statistics & Data Analysis*, 56, 2303–2316. MR2897581
[8] Chen, Z., Tian, Z. and Zhao, C. (2012c). Monitoring persistence change in infinite variance observations. *Journal of the Korean Statistical Society*, 41, 61–73. MR2933215
[9] Chen, Z. and Tian, Z. (2014). Ratio test for variance change in nonparametric regression. *Statistics*, 48(1), 1–16. MR3175754
[10] Chu, C. S. J., Stinchcimbe, M. and White, H. (1996). Monitoring structural change. *Econometrica*, 64, 1045–1065.

[11] Davidson, J. and De Jong, R. M. (2000). The functional central limit theorem and weak convergence to stochastic integrals II: fractionally integrated processes. *Econometric Theory*, 16(5), 643–666. MR1802836
[12] Davidson, J. and Hashimzade, N. (2009). Type I and type II fractional Brownian motions: a reconsideration. *Computational Statistics & Data Analysis*, 53(6), 2089–2106. MR2665097
[13] De Long, J. B. and Summers, L. B. (1988). How does macroeconomic policy affect output? *Brookings Papers on Economic Activity*, 2, 433–494.
[14] Elliott, G., Rothenberg, T. J. and Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64, 813–836. MR1399219
[15] Gerolimetto, M. and Procidano, I. (2008). A test for fractional cointegration using the sieve bootstrap. *Statistical Methods & Applications*, 17, 373–391. MR2425192
[16] Gombay, E. and Serban, D. (2009). Monitoring parameter change in AR(p) time series models. *Journal of Multivariate Analysis*, 100, 715–725. MR2478193
[17] Hassler, U. and Scheithauer, J. (2011). Detecting changes from short to long memory. *Statistical Papers*, 52, 847–870. MR2846689
[18] Hassler, U. and Meller, B. (2013). Detecting multiple breaks in long memory: the case of U.S. inflation. *Empire Econmics*. DOI: 10.1007/s00181-013-0691-8.
[19] Kapetanios, G. and Psaradakis, Z. (2006). Sieve bootstrap for strongly dependent stationary processes. Working paper no. 552, Queen Mary, Department of Economics, University of London.
[20] Kejriwal, M., Perron, P. and Zhou, J. (2012). Wald tests for detecting multiple structural changes in persistence. *Econometric Theory*, 29(2), 289–323. MR3042757
[21] Kim, J. Y. (2000). Detection of change in persistence of a linear time series. *Journal of Econometrics*, 95, 97–116. MR1746619
[22] Kirch, C. (2008). Bootstrapping sequentical change-point tests. *Sequential Analysis*, 27, 330–349. MR2446906
[23] Leybourne, S. J., Taylor, R. and Kim, T. H. (2007a). CUSUM of squares-based tests for a change in persistence. *Journal of Time Series Analysis*, 28, 408–433. MR2355315
[24] Leybourne, S. J., Kim, T. H. and Taylor, A. M. B. (2007b). Detecting multiple changes in persistence. *Studies in Nonlinear Dynamic & Economics*, 11, 1370–1384.
[25] Martins, L. F. and Rodrigues, P. M. M. (2012). Testing for persistence change in fractionally integrated models: an application to world inflation rates. *Computational Statistics & Data Analysis*. DOI:10.1016/j.bbr.2011.03.031.
[26] Park, J. Y. (2002). An invariance principle for sieve bootstrap in time series. *Economic Theory*, 18, 469–490. MR1891831
[27] Poskitt, D. S. (2008). Properties of the sieve Bootstrap for fractionally integrated and non-invertible processes. *Journal of Time Series Analysis*, 29(2), 224–250. MR2423112
[28] Sibbertsen, P. and Kruse, R. (2009). Testing for a break in persistence under long-range dependences. *Journal of Time Series Analysis*, 30(3), 263–285. MR2512240
[29] Steland, A. (2006). A bootstrap view on Dickey-fuller control charts for AR(1) series. *Austrian Journal of Statistics*, 35, 339–346.
[30] Steland, A. (2007). Monitoring procedures to detect unit roots and stationarity. *Economic Theory*, 23, 1108–1135. MR2397197

Zhanshou Chen
Department of Mathematics
Qinghai Normal University
Xining, Qinghai 810008
P.R. China
E-mail address: chenzhanshou@126.com

Zheng Tian
Department of Applied Mathematics
Northwestern Polytechnical University
Xi'an, Shaanxi 710072
P.R. China
E-mail address: czshhh@mail.nwpu.edu.cn

Yuhong Xing
Department of Mathematics
Qinghai Normal University
Xining, Qinghai 810008
P.R. China
E-mail address: 602416983@qq.com